DOCUMENT RESUME

ED 105 336 95 CG 009 706

AUTHOR	Krus, David J.
TITLE	Synopsis of Basic Theory and Techniques of Order Analysis. Occasional Paper No. 26.
INSTITUTION	Minnesota Univ., Minneapolis. Research, Development, and Demonstration Center in Education of Handicapped Children.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
REPORT NO	Occas-Pap-26
BUREAU NO	332189
PUB DATE	Apr 74
GRANT	OEG-09-332189-4533(032)
NOTE	163p.; For related document, see CG 009 707
EDRS PRICE	MF-$0.76 HC-$8.24 PLUS POSTAGE
DESCRIPTORS	*Information Theory; Mathematical Logic; Matrices; *Measurement Techniques; *Models; Research Projects; *Statistical Analysis; *Theories; Validity
IDENTIFIERS	*Order Analysis

ABSTRACT

Order analysis is a measurement technique for the quantitative and generalizable description of complex structures. Relations between the manifest elements of a data matrix and its underlying latent entities and attributes are considered in terms of magnitudes as basic scale units. Magnitude is conceptualized as a difference relation, translated into the binary units of information theory and linked with the functional relations of the propositional calculus of formal logic. Deterministic and probabilistic models of order analysis are developed. The deterministic model elaborates an algorithm for partitioning of binary relations within the context of Guttman-type scaling. The probabilistic model extends these principles within the framework of graph theory; it permits the generalization of ordered structures into their latent domains with a known degree of statistical certainty. Relation of order analysis to factor analysis is discussed, as well as components of isolated binary variation. Order analysis is validated on sets of prestructured, distorted and random data and its utility considered within the general framework of multivariate analysis models. (Author)

ED105336

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to
CG
EC
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

OCCASIONAL PAPER #26

Project No. 332189
Grant No. OE-09-332189-4533 (032)

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

BEST COPY AVAILABLE

Synopsis of Basic Theory and Techniques of Order Analysis

David J. Krus
University of Minnesota
Research, Development and Demonstration
Center in Education of Handicapped Children
Minneapolis, Minnesota

April 1974

CG 009 706

Department of Health, Education and Welfare

U. S. Office of Education

Bureau of Education for the Handicapped

TECHNICAL REPORTS

University of Minnesota Research, Development and Demonstration
Center in Education of Handicapped Children

(Place of publication shown in parentheses where applicable)

1. D. Krus. Synopsis of basic theory and techniques of order analysis. Occasional Paper #26. April 1974.
2. S. Samuels, J. Spiroff & H. Singer. Effect of pictures and contextual conditions on learning to read. Occasional Paper #23. March 1974.
3. A. Taylor, M. Thurlow & J. Turnure. Elaboration as an instructional technique in the vocabulary development of EMR children. Research Report #59. March 1974.
4. N. Buium & J. Turnure. The universality of self-generated verbal mediators as a means of enhancing memory processes. Research Report #58. January 1974.
5. D. Moores, K. Weiss, & M. Goodwin. Evaluation of programs for hearing impaired children: Report of 1972-73. Research Report #57. December 1973.
6. J. Turnure, W. Charlesworth, D. Moores, J. Rynders, M. Horrobin, S. Samuels, & R. Wozniak. American Psychological Association Symposium Papers. Occasional Paper #24. December 1973.
7. N. Buium. Interrogative types of parental speech to language learning children: a linguistic universal? Research Report #56. December 1973.
8. D. Krus. An outline of the basic concepts of order analysis. Occasional Paper #23. February, 1974.
9. D. Krus. Order analysis: A fortran program for generalizable multidimensional analysis of binary data matrices. Occasional Paper #22. November 1973.
10. W. Bart. The pseudo-problem of IQ. Occasional Paper #21. October 1973.
11. J. Turnure & M. Thurlow. Verbal elaboration and the enhancement of language abilities in the mentally retarded: The role of interrogative sentence-forms. Occasional Paper #20. October 1973.
12. P. Dahl, S. Samuels & T. Archwamety. A master based experimental program for teaching poor readers high speech word recognition skills. Research Report #55. September 1973.
13. R. Riegel, F. Danner & L. Donnelly. Development trends in the generation and utilization of associative relations for recall by EMR and non-retarded children: The SORTS test. Research Report #54. August, 1973.
14. R. Hoffmeister & D. Moores. The acquisition of specific reference in the linguistic system of a deaf child of deaf parents. Research Report #53. August 1973.
15. W. Bart & M. Smith. An interpretive framework of cognitive structures. Occasional Paper #19. June 1973.
16. C. Clark & J. Greco. MELDS (Minnesota Early Language Development Sequence) Glossary of Rebuses and Signs. Occasional Paper #18. June 1973.
17. J. Turnure. Interrelations of orienting response, response latency and stimulus choice in children's learning. Research Report #52. May 1972.
18. S. Samuels & P. Dahl. Automaticity, reading and mental retardation. Occasional Paper #17. May 1973.
19. S. Samuels & P. Dahl. Relationships among IQ, learning ability, and reading achievement. Occasional Paper #16. May 1973.
20. N. Buium & J. Rynders. The early maternal linguistic environment of normal and Down's Syndrome (Mongoloid) language learning children. Research Report #51. May 1973.
21. T. Archwamety & S. Samuels. A mastery based experimental program for teaching mentally retarded children word recognition and reading comprehension skills through use of hypothesis/test procedures. Research Report #50. May 1973.
22. W. Bart. The process of cognitive structure complexification. Research Report #49. April 1973.
23. B. Best. Classificatory development in deaf children: Research on language and cognitive development. Occasional Paper #15. April 1973.
24. R. Riegel, A. Taylor, & F. Danner. The effects of training in the use of a grouping strategy on the learning and memory capabilities of young EMR children. Research Report #48. April 1973.
25. J. Turnure & M. Thurlow. The latency of forward and backward association responses in an elaboration task. Research Report #47. March 1973.
26. R. Riegel & A. Taylor. Strategies in the classroom: A summer remedial program for young handicapped children. Occasional Paper #14. March 1973.
27. D. Moores. Early childhood special education for the hearing impaired. Occasional Paper #13. February 1973.
28. R. Riegel & A. Taylor. A comparison of conceptual strategies for grouping and remembering employed by educable mentally retarded and non-retarded children. Research Report #46. February 1973.
29. J. Rynders. Two basic considerations in utilizing mothers as tutors of their very young retarded or potentially retarded children. Occasional Paper #12. January 1973.
30. R. Bruininks, J. Rynders & J. Gross. Social acceptance of mildly retarded pupils in resource rooms and regular classes. Research Report #45. January 1973.

ED105336

Occasional Paper # 26

Project No. 332189
Grant No. OE-09-332189-4533 (032)

Synopsis of Basic Theory and Techniques of Order Analysis

David J. Krus
University of Minnesota
Research, Development and Demonstration
Center in Education of Handicapped Children
Minneapolis, Minnesota

April 1974

The research report herein was performed pursuant to a grant from the Bureau of Education for the Handicapped, U.S. Office of Education, Department of Health, Education, and Welfare to the Center for Research and Development in Education of Handicapped Children, Department of Special Education, University of Minnesota. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgement in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.

Department of Health, Education, and Welfare
U. S. Office of Education
Bureau of Education for the Handicapped

RESEARCH AND DEVELOPMENT CENTER
IN EDUCATION OF HANDICAPPED CHILDREN
Department of Special Education
Pattee Hall, University of Minnesota, Minneapolis, Minnesota 55455

The University of Minnesota Research, Development and Demonstration Center in Education of Handicapped Children has been established to concentrate on intervention strategies and materials which develop and improve language and communication skills in young handicapped children.

The long term objective of the Center is to improve the language and communication abilities of handicapped children by means of identification of linguistically and potentially linguistically handicapped children, development and evaluation of intervention strategies with young handicapped children and dissemination of findings and products of benefit to young handicapped children.

Preface

Identification of unidimensional sequences and hierarchical orders is a fundamental task of measurement theory. In the area of education, proper sequencing and gradation of learning experience often determines the overall success or failure of educational intervention. The problem of adequate determination of a particular degree of cognitive development and subsequent interfacing of this intellectual "readiness" with properly scaled sequences of educational process, is salient in the area of special education.

As explicated later, extant psychometric methods, traditionally applied to the problem of isolation of ordered, hierarchically graded, unidimensional components of data matrices, are not quite satisfactory. The initial effort to reconceptualize the traditional approach in terms of a logic model was done in collaboration with Dr. William Bart. Combining the earlier findings of Dr. Peter Airasian of Boston College with some new insights, the resulting procedure for isolation of logical orders among variables was called "tree" and later "ordering" theory.

Results of practical applications of ordering theory methods were only partially encouraging. The formal, logic-based orderings led to meaningful interpretations only in special cases when the property measured possessed a singular meaning. Also, the inferential rationale was nonexistent and attempts to construct logic diagrams of any sizable data matrices resembled in favorable cases an intricate maze. In the case of medium-sized, multidimensional data matrices typical of most research problems, the logic diagram was too complicated to be interpreted.

The method of order analysis was initially concieved through reconceptualization of the traditional variance measure, by identification of functions of propositional calculus which would be both information and variance generative, and by realization that separate branches of "tree" logic diagrams of ordering theory are identical to Guttman-type data matrices. The subsequent development of order analysis was documented in a series of research reports, occassional papers, journal articles, and yet unpublished manuscripts. It is presented here in an integrated form with the exception of the Fortran program for order analysis, published separately in the same series of technical reports.

The present synopsis of research findings spans the development of order analysis from the spring of 1970. During the intervening years, this development was supported by the Eva O. Miller Reserach Fellowship (1972-1973), by a grant of computer time from the University Computer Center (1971-1974) and by a grant to the Research, Development and Demonstration Center (1973-1974). The Center is funded by a grant (OE-09-332189-4533-032) from the United States Office of Education.

I am indebted to Patricia Bland for editorial assistance, to Drs. David Weiss and William Bart for critical comments on the manuscript and to Dr. Donald Moores for making the preparation of this report possible.

David J. Krus

Minneapolis, Minnesota
April, 1974

Abstract

Order analysis is a measurement technique for the quantitative and generalizable description of complex structures. Relations between the manifest elements of a data matrix and its underlying latent entities and attributes are considered in terms of magnitudes as basic scale units. Magnitude is conceptualized as a difference relation, translated into the binary units of information theory and linked with the functional relations of the propositional calculus of formal logic. Deterministic and probabilistic models of order analysis are developed. The deterministic model elaborates an algorithm for partitioning of binary relations within the context of Guttman-type scaling. The probabilistic model extends these principles within the framework of graph theory; it permits the generalization of ordered structures into their latent domains with a known degree of statistical certainty. Relation of order analysis to factor analysis is discussed, as well as components of isolated binary variation. Order analysis is validated on sets of prestructured, distorted and random data and its utility considered within the general framework of multivariate analysis models.

CONTENTS

preface ii
abstract iv

CHAPTER

1: INTRODUCTION General polynomial model of test theory. Assumption of linear orders. Traceability of equivalence of inceptive and terminal structures. Generalizability of formal structures and related problems. 1

2: THE DIFFERENCE RELATION Magnitudes as a difference relation. Difference relation and measurement in psychology. Generation of the difference relation. Difference relation, information and variance. 15

3: THE RELATIONS BETWEEN COGNITIVE, LOGICAL AND GEOMETRIC STRUCTURES Cognitive and logical structures. Dimensionality and order. Logical and geometric structures. Conceptual relations between order and factor analyses. 37

4: THE ELEMENTS OF ORDER ANALYSIS The prerequisite and disconfirmatory relationships. The logic diagram of a test space (the nonmetric model A). The logic diagram of a unit space model B. Dimensionality and reproducibility. 57

5: THE DETERMINISTIC MODEL OF ORDER ANALYSIS An outline of major computational steps. Additive solution for the data matrix dimensionality. 75

6: THE PROBABILISTIC MODEL OF ORDER ANALYSIS Probabilistic structure of Guttman-type data matrices. Unidimensional components of the connected graphs and their extraction. Latent dimensions of the data matrix. 92

7: VALIDITY OF ORDER ANALYSIS Prestructured data (monotonic transformations). Distorted data (nonmonotonic transformations). Random data. 117

8: RETROSPECT AND PROSPECTUS Standardization. Optimization of the dominance-consonance ratios. Thurstone's Law of Comparative Judgement. Types of variance extracted. Integrated analysis of the general relational space. 1[illegible]

REFERENCES 144

BEST COPY AVAILABLE

Chapter 1

INTRODUCTION

The foundations of measurement in psychology formulate basic axioms and theorems pertaining to the question of how abstract properties can be expressed within abstract systems of constructs. This lack of tangible referents is characteristic of measurement in social sciences as contrasted with its collateral branch in the physical sciences.

Measurement implies the description of data in terms of numbers, and any attempt to build a structure of a collection of data usually centers on the possible operations which can be done on such a set of numbers. It is difficult to imagine any kind of measurement which cannot be expressed in a matrix or vector form. From the standpoint of numerical analysis it is irrelevant how elements of a data matrix were obtained. Also the content of measurement, the domain of attributes or entities inferred or assumed is totally irrelevant with respect to the mathematical operations on the elements of the data matrix.

On the other hand, research in science is dependent jointly on the validity of measurement models and the validity of the data; both being necessary but not sufficient conditions for the validity of the total measurement operation.

Considered on a general level, operations on which a system of measurements of complex systems is founded can be unfolded in two basic directions. The more familiar direction is toward gradually increasing complexity. The other direction, which is less familiar, leads to abstractness and logical simplicity. It endeavors to extend operational systems by scrutinizing their general principles and assumptions, and attempts to redefine them at a more basic level. Subsequent development can then reintroduce a reconstruction of formerly primitive concepts and continue in their elaboration along alternative paths.

In noncontrived situations, measurement is logically rather an intricate operation. In the case of nonsingular systems, its complexity is based on the fact that particular properties contained in the system are initially unknown. The matrix of numbers should have a definite meaning and not only certain formal properties. In order for this to be true, the relation of isomorphism must be established between the latent properties of the natural system and the formal one. To preserve this isomorphism during often elaborate transformations, the first step in the analysis should aim at identification of these attributes. By establishing a proper isomorphic relation between these attributes and our formal, quantitative, numerical system, any lawful algebraic operation will hold and will have a definite, not only formal, meaning.

The question of validity of measurement operations cannot be treated in its complexity without considering jointly the formal properties of quantitative systems and the natural properties of systems capable of processing information, which are frequently described as the cognitive systems. The problem is phrased here as the inquiry into the joint interaction of both formal (numerical and logical) and natural (cognitive) systems and contrasted with measurement operations based on the general polynomial model of the test theory (Horst, 1966).

General polynomial model of test theory.

Classical test (measurement) theory may be briefly described as an application of linear least squares estimation along the lines of traditional error analysis and is perhaps best typified by its major compendia (Yule, 1919; Kelley, 1924; Spearman, 1926; Guilford, 1936; Thurstone, 1947; Gulliksen, 1950; Torgerson, 1958; Lord and Novick, 1968). While some models of this group compel partitioning of a data matrix into submatrices and use solutions of simultaneous normal equations for determination of the transformation matrix (as in the case of multiple regression analysis), other models transform the whole data matrix according to a specific algorithm. A typical example of this latter case is the factor analytic model. It consists of a chain of matrix transformations of the data matrix, the key transformation being the solution of the characteristic equation for its roots or factors. This particular transformation attempts to answer the question whether there is a scalar number λ which when multiplying the vector $\underline{x}$, yields the identical vector resulting from the product $\lambda\underline{x}$. The polynomial resulting from

$$\det(A-\lambda I) \qquad (1.1)$$

is called the characteristic polynomial. In order to find its roots, we set it equal to zero, thus changing it into the characteristic equation. Its roots, or factors are sometimes called eigenvalues and can be imagined as intersection points of mutually orthogonal axes with a function specified by the characteristic equation. In general it can be shown that the characteristic equation of a square matrix of order n equals a polynomial of nth order:

$$\det(A-\lambda I) = {}^{n}+c_{n-1}{}^{n-1}+c_{n-2}{}^{n-2}+\ldots+c_{1+(-1)^{n}[A]} \qquad (1.2)$$

Comparison of this polynomial with another model of the classic measurement theory as e.g. a model of a typical ANOVA design including interaction terms

$$y = w_1x_1+w_2x_2+w_3x_1x_2\ldots \qquad (1.3)$$

or a multiple regression model

$$y = w_1x_1+w_2x_2+w_3x_3\ldots \qquad (1.4)$$

allows for tentative postulation of a nonspecific general multivariate polynomial model. Suggestion of this general paradigm was made by Horst (1966), and we can observe most, if not all of the classical theory models to be subsummed by it. Its integrative potential is suggested e.g. by Burt's (1966) conjoint model for both factor analysis and the analysis of variance.

Within the framework of classic measurement theory, development can be observed from zero-order, partial, and multiple correlations into the region of multiple regression, component analysis, canonical correlation, factor analysis, and related techniques. Relatively recent developments attempt to extend the general multivariate polynomial model to higher order or non-linear functions as in the case of the recruitment equations of latent structure analysis

$$P_{ijk} + a^{0}_{ijk} + a^{1}_{ijk}M_1 + a^{2}_{ijk}M_2 + a^{3}_{ijk}M_3 \qquad (1.5)$$

(Lazarsfeld, 1958; Gibson, 1960, 1966).

Lunneborg (1960) suggests the existence of a close relationship between the configural analysis (McQuitty, 1955) and the latent structure models. Horst, in a similar context, comments that "it can readily be shown that the general multivariate polynomial model will completely identify all possible patterns of item responses to a set of items. The general configural model may then be taken as the general multivariate model considered in latent class analysis." (Horst, 1966, p. 150).

Psychometric theory connects polynomial models into a fairly integrated whole. Inherent in this group of models are certain difficulties which by their nature and pervasiveness generate doubts if they can be solved without changing the basic concepts inherent in the general polynomial paradigm. These difficulties are highlighted as we move from bivariate to multivariate polynomial models and are best typified by perhaps the most sophisticated model of the polynomial class, represented by the factor analytic group.

Assumption of linear orders

The majority of transformation matrices within the classic model were developed within the linear algebra context of methods allowing for simultaneous solutions of systems of linear equations. Thus e.g. Spearman's (1904) criterion of tetrad differences as well as Yule's (1919) phi coefficient are applications of the principle of the determinant. The assumption of the presence of a linear order among the elements of a data matrix is seldom tested for, and "orderings developed from both logical and statistical analysis indicate that non-linear orderings are the rule rather than the exception" (Airaisian and Bart, 1973). Failure to take into consideration interitem relations of different types can result in substantial loss of information.

Meehl (1950) illustrated with theoretical data that two test items taken jointly as a pattern or response can have a perfect correlation with a criterion even though, when treated separately, this correlation vanishes. This obviously represents the extreme case which, however, can exist in lesser degree. It is interesting to compare tables 1, 2, and 3 of Meehl's original article (Meehl, 1950, p. 166) with Table 1.1. Meehl's fourfold point surfaces were translated into the original data matrix form and responses of one hundred subjects were reduced into four subject types. Applying the method described by Krus and Bart (1974), the evident bidimensionality of this matrix can be verified. Meehl's paradox thus exemplifies the need for priority of dimensional decomposition of a data matrix over other computational operations.

Attempts to circumvent problems highlighted by Meehl's paradox lead to methods typified by Lykken's (1956) actuarial pattern analysis and McQuitty's (1957) multiple configural analysis. This approach stresses a typological as opposed to dimensional way of organization of test responses. It encounters the basic difficulty that the number of response patterns increase exponentially with the number of items. This led to difficulties such as the "problem of empty cells," which were

Table 1.1. Reconstruction of the Data Matrix for the Meehl Paradox. Table 1.1 is the rearranged composite of Meehl's Tables 1, 2, and 3 (Meehl, 1950, p. 166).

SUBJECT TYPES	ITEMS	DIAGNOSIS
SUBJECT TYPE A	1 1	NEUROTIC
SUBJECT TYPE B	0 0	NEUROTIC
SUBJECT TYPE C	1 0	SCHIZOPHRENIC
SUBJECT TYPE D	0 1	SCHIZOPHRENIC

yet not solved.

Similar problems are encountered when non-linear relationships are assumed and a higher order multivariate polynomial model is used. Limitation of such a procedure is described by Horst as "the enormous rapidity with which higher-order variables are introduced into the system [which] frequently causes one to run out of degrees of freedom". (Horst, 1966, p. 135)

These and related findings point toward the possible loss of information induced by this assumption; the amount of this loss is determined by the character of the data analyzed.

Traceability of equivalence of inceptive and terminal structures.

Construction of isomorphic and homomorphic formal structures of biological systems (especially human cognitive systems) is the central area of psychological inquiry. Isomorphism means similarity in pattern, where the relations between the domain and range of mapping operations are undistorted. Thus a photographic negative of a chessboard will correctly reflect spatial relations between pieces, though the appearances with respect to brightness are different. Homomorphism occurs when (isomorphic) a "many-one" type of transformation is applied to a system. Problems of homomorphic transformations are traditionally treated (in psychology) in connection with problems of data types and their corresponding scale types. The matter in question is the relation between the character of property contained by the data matrix and the character of numbers intended as its symbolic representation. Following the writings of Stevens (1951), the fifties witnessed the period of interest of psychologists in the theory of numbers and their role in measurement, which culminated in Torgerson's (1958) comprehensive treatment of the subject. Since then interest in the whole problem declined, as evidenced by the cursory treatment of this topic by Lord and Novick (1968, pages 20-23). This gradual decline was paralleled at the same time by the shift of attention from the area of finite mathematics to the treatment of measurement as continuous. This is contrary to the development in other social sciences, where the computer revolution was parallelled by the rise of finite mathematics (cf. Kemeny et al., 1972).

Stevens' treatment of the problem of the "assignment of numerals to objects or events" (1951, p. 22) was stimulated by the earlier transfer of the topic from mathematics into philosophy (cf. Campbell, 1928, 1938; Carnap, 1950; Russell, 1903). The original inquiry into the basic problems of the theory of numbers is by Frege (1884, 1893). Recent theoretical works of Adams et al. (1965), Luce and Tukey (1964), and Suppes and Zinnes (1963) elaborate on correspondence between the measured properties and their numerical representations.

The requirement that the measured property should be represented by a symbolic system based minimally on a homomorphic transformation is reasonable. Consider the following homomorphic transformation:

$$T_1: \Big\downarrow \begin{array}{ccccccccc} a & b & c & d & e & f & g & h & i \\ 1 & 1 & 1 & 1 & 2 & 2 & 3 & 3 & 3 \end{array} \tag{1.6}$$

This transformation of the more complex system into a simpler one insures that subsequent operations carried on the simpler system will be based on a subset of information contained by the original system.

Assumptions of some recent measurement models, derived within the classical framework, try to superimpose a stronger model on weaker data. Thus Lord and Novick state that "we shall treat a measurement as having interval scale properties, although it is clear that the measurement procedure and the theory underlying it yield only a nominal, or, at best an ordinal scale" (Lord and Novick, 1968, p. 22). The application of strong models on weak data amounts to a converse of homomorphic transformation. This produces possible distortion of results by model-generated noise and difficulty in tracing properties of the original system within the derived model. The presence of these distorting factors was felt by the same authors, who ex post facto rather paradoxically value utility over validity: "In fact, at least with psychological models of the type considered here, it can be taken for granted that every model is false [sic] and that we can prove so, if only we collect a sufficiently large sample of data. The key question, then, is the practical utility of the model, and not its ultimate truthfulness." (Lord and Novick, 1968, p. 383).

If the properties measured are initially unknown, the relation between them and the corresponding formal system is established in two ways. First, the empirical approximation of an isomorphic relation can be achieved a posteriori of the initial operations aimed at their extraction and definition. The second way is the a priori definition of the property measured, an approach frequently adopted within the context of latent test theory (Lord, 1952) and related models (Rasch, 1960). The typical definition of the a priori type is: "the trait or ability under discussion can be thought of as an ordered variable represented numerically in a single dimension" (Lord, 1966, p. 22). In noncontrived situations, the validity of this "a priori approach" is doubtful. In numerous instances of "real" data matrices, isomorphic relation between properties present in the data matrix and derived structures is distorted and untraceable to an unknown degree by this a priori definition of the measured property. Consequently, one of the conditions necessary to establish a valid relation between inceptive and terminal structure is an early attempt to isolate properties we intend to measure. This takes priority over any other goals of the analysis.

It is interesting to note that criticism of this tendency to disregard possible (and probable) multidimensional structure of data (i.e. the presence of multiple properties in the data matrix) was made as early as 1843: "When determination of the mean is applied to the different parts of a complicated system, it should be carefully kept in mind that these mean values might be inconsistent with each other, for the system might be in an impossible state if each of the elements took on its mean value, determined separately." (Cournot, 1843, p. 213).

Generalizability of formal structures and related problems.

As contrasted with the assumption of linear orders, the problem of generalizability was frequently explicitly stated and its solutions were attempted both within and outside of classical test theory. This problem was usually classified together with other long-lasting issues of factor analysis such as problems of communality estimation, and number of factors to extract or retain. (cf. Thurstone, 1947; Harman, 1960; Horst, 1965; Rummel, 1970; Weiss, 1970, 1971) In this context, the problem of factorial invariance was often blamed for the multiplicity of isolated structures and consequent proliferation of theoretical constructs. The well known example is the case of constructs associated with the MMPI item pool (Wiggins, 1968), or the case of Cattel's (1957) and Guilford's (1943) factorially derived traits.

The instability of factorial structures is a direct consequence of a substantial lack of statistical inferential rationale inherent in present models. This fact is further complicated by the well-known tendency of these models to isolate patterns of random variation. This was perhaps best illustrated when Armstrong and Soelberg (1968) used principal components analysis to analyze parts of Rand Corporation tables of random numbers, where arbitrary trait names were assigned prior to the analysis and convincingly interpreted afterwards.

Effort from within factor analytic theory to remedy these problems include Bartlett's (1950) test of significance of the correlation matrix prior to factor analysis through Humphreys and Ilgen's (1969) "parallel analysis" approach. Attempts to build the inferential mechanism directly into the model itself include Kaiser and Caffrey's (1965) alpha factor analysis, generalizing from a sample of variables to its population, and Rao's (1955) canonical factor analysis, generalizing to the population of subjects.

Harris' (1962) comparative theoretical study of the mathematical relationships between some traditional models showed the proportionality between their factorial structures as dependent on the rescaling of the diagonal of the correlation matrix, resulting in varying degrees of depression of its rank. In general, "one of the main characteristics that differentiate one method of factoring from another is the procedure used for estimating the communalities" (Nie, Bent, and Hull, 1970, p. 212). Thorndike's (1970) empirical comparative study of structures returned by five different factor analytic models (principal components, minimum residual, maximum likelihood, image, and alpha), illustrates the consequences of these alterations of the main diagonal. The same data were described from 2 to 11 dimensions depending on which method was employed. In the case of non-psychological data matrices and methods other than principal components, failure to obtain a solution frequently occurred.

Cluster analytic and multidimensional scaling models represent the main alternatives to factor analysis. Borgen and Weiss' (1971) overview of replicability and validity of a variety of cluster analytic methods reports that results of cluster analytic methods do not differ radically from those achieved by factor analytic methods. These authors conclude that "probably one of the most serious hidden dangers of cluster analysis is the possibility of basing a cluster analysis completely or primarily on random data. With most clustering methods it is entirely possible to perform a superficially normal cluster analysis on totally random data." (Borgen and Weiss, 1971, p. 590). Coombs' description of the generalizability of multidimensional scaling methods [represented by the work of Shepard (1962), Kruskal (1964), Torgerson (1965); Lingoes (1966), Guttman (1968), and Beals, Krantz, and Tversky (1968)] concludes that "statistical theory for such analyses is almost totally undeveloped" (Coombs, Dawes, and Tversky, 1970, p. 76). Coombs' judgment is in concordance with Cliff's recent admission: "Among the unfilled needs is some sort of statistical basis for deciding such questions as the number of dimensions which can be readily defined and the uncertainty of the coordinates of a given point. This would be much more valuable than tests of the null (nullest!) hypothesis that the proximities are random. Failing such developments, multidimensional scaling is likely to drift into the inferential quagmire that holds most of traditional factor analysis" (Cliff, 1973, p. 484).

Recent advances in computer technology, facilitating applications of theoretical developments in the general area of finite mathematics, allow for workable alternatives to former models of dimensional analysis. Order analysis attempts to introduce a sound inferential rationale into this area; the model is based on fundamental logical processes, which increases the probability of preservation of isomorphism between the initial and derived structures. It does not assume the presence of linear orders in the data matrix and attempts to generalize the sample spaces of attributes and entities into their respective domains.

Chapter 2

THE DIFFERENCE RELATION

Magnitude as a difference relation.

In the typical measurement situation in psychology, a group of subjects is presented with a set of statements and their reactions to these statements are recorded in the data matrix. Thus created inceptive formal structure is transformed and resulting terminal structure interpreted. This repetitive process of scientific inference is based on certain primitive notions. Variance, defined as a second moment about the mean, is frequently considered a primitive concept with respect to partitioning of variance by various psychometric models. In order analysis, the concept of magnitude as a difference relation is used as a basic measurement unit.

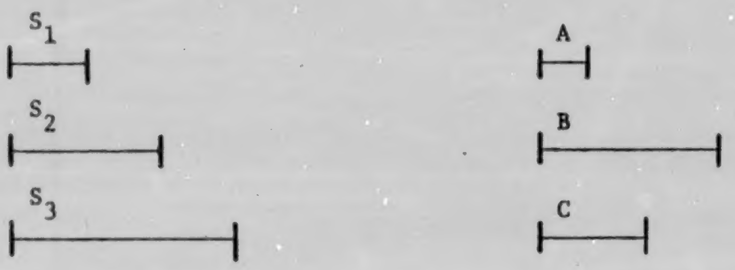
The concept of magnitude has definite philosophical underpinnings. It was developed by Russell (1903) within the context of the Kantian epistemological controversy, centering around the nature of meaning of true propositions of arithmetic. To substantiate his argument that the context of formal structures is determined by truth functions of formal logic, Russell had to refute one of Kant's axioms of intuition, namely the axiom of anticipation of perception. This axiom states that "every reality in phenomena, however small it may be, has a degree, that is, an intensive quantity which can always be diminished." (Kant, T. M. Greene edition, 1929, p. 115).

To support his argument, Russell had to distinguish between quantity and magnitude. He defines magnitude as "anything which is greater or less than something else" (Russell, 1903, p. 159) and rejects the notion of direct report of quantity as equality of a unit as not definable in terms of logical constants and therefore, "not properly a notion belonging to pure mathematics at all." (Russell, 1903, p. 158).

The importance of the distinction between magnitude and quantity was not recognized, although frequently reported in psychometric literature (Guilford, 1954, p. 7; Torgerson, 1958, p. 26). In the typical example of length, for any two physical objects the successive laying off of the second object some finite number of times results in the second object going farther than the first. Popular notion assumes (as early Greek mathematicians did), that the resulting length could be reported as whole numbers only if the units were made small enough. This assumption of equality of minute units can be disproved on the basis of the Pythagorean theorem. Accepting this to be true, the actual length of observed differences (when compared with some standard) is then its magnitude.

Consider a case of lines compared by a set of standards as shown in Table 2.1. Dimensionality is given by a definition

Table 2.1. Computation of magnitudes of a set of three lines A, B, C as compared with metric-system based standards. S_1=1cm, S_2=2cm, and S_3=3cm.



$S_1A=0$ $S_2A=0$ $S_3A=0$

$S_1B=1$ $S_2B=1$ $S_3B=0$

$S_1C=1$ $S_2C=0$ $S_3C=0$

Table 2.2 Ordered magnitudes of lines A, B, C as compared with a set of metric standards S_1, S_2, S_3 from Table 2.1.

		LINES		
		B	C	A
STANDARDS	S_1	1	1	0
	S_2	1	0	0
	S_3	0	0	0
MAGNITUDES		2	1	0

of a line, magnitude is reported as 1 if line is greater than a particular standard, and as 0 if line is smaller than a standard. By reordering lines according to their increasing magnitudes, a Guttman-type data matrix (Guttman, 1941) can be obtained, as shown in Table 2.2. Magnitudes of lines drawn in Table 2.1 are 2, 1, and 0 for lines B, C, and A respectively. Quantities of these lines can be reported as .5 cm (A), 2.5 cm (B) and 1.5 cm (C).

In the measurement of length, this principle of comparison with a standard is rather straightforward and is accomplished by the direct observation of unit concatenation in one-dimensional serial order. Both order and dimensionality therefore need not be expressed in logical terms; they are overtly present and covertly assumed. This is usually not true in the case of psychological measurement, where both order and dimensionality are frequently initially unknown.

Difference relation and measurement in psychology.

The significance of the difference relation for measurement in psychology was recognized by Fechner (1871), who introduced the concept of "juxtapositions of cognitions" as a basic unit of the method of pair comparisons. As an introduction into the scope and problems of this method let us consider the hypothetical reconstruction of steps taken by Friedrich Moh in 1820 to

FRAME 2.1. THE N=1 PAIR COMPARISONS EXPERIMENT.

Below is a series of statements pertaining to the hardness of minerals. Please record your observations as "yes" or "no" regarding the mineral surface changes.

Is diamond scratched by topaz?
Is quartz scratched by topaz?
Is gypsum scratched by topaz?
Is topaz scratched by diamond?
Is quartz scratched by diamond?
Is gypsum scratched by diamond?
Is topaz scratched by quartz?
Is diamond scratched by quartz?
Is gypsum scratched by quartz?
Is topaz scratched by gypsum?
Is diamond scratched by gypsum?
Is quartz scratched by gypsum?

Table 2.3. The data matrix.

	T	D	Q	G	Σ
T	0	0	1	1	2
D	1	0	1	1	3
Q	0	0	0	1	1
G	0	0	0	0	0

Table 2.4. Reordered data matrix.

	D	T	Q	G	Σ
D	0	1	1	1	3
T	0	0	1	1	2
Q	0	0	0	1	1
G	0	0	0	0	0
Σ	0	1	2	3	6

develop his scale of hardness of minerals. His measurement was based on the direct observations of results of the "scratching" operation. The domain sampled was the class of all minerals. The property measured was conceptualized as hardness and was directly accessible to his empirical observations.

Imagine that Friedrich Moh's original sample consisted of gypsum, quartz, topaz, and diamond (G, Q, T, D), and that he was at the same time the subject and observer in this N=1 pair comparisons experiment. This experiment could be arranged as shown in Frame 2.1. It may be observed that nothing is assumed but the ability to observe a change and record its occurrence. The data matrix for this experiment can be constructed if we adopt a convention that "YES" will be assigned the value 1, and "NO" the value of 0 if the change observed was in row-column direction (Table 2.3). The meaning of the elements of this data matrix is direct. Three changes of surfaces were observed when scratching was done with a diamond, two when with topaz, etc.

Elements of both the data matrix (Table 2.3) and the reordered data matrix (Table 2.4) are magnitudes. Table 2.4 was reordered according to decreasing row and increasing column marginals, i.e. in such a way that marginal apex will be located in the lower right corner and the matrix apex in

FRAME 2.2. THE RATING SCALE EXPERIMENT

Below is a series of the names of various minerals accompanied by rating scales. Please indicate the extent to which you feel every mineral could be characterized as hard or soft.

TOPAZ

soft:_:_:_:X:_:hard

DIAMOND

soft:_:_:_:_:X:hard

QUARTZ

soft:_:_:_:X:_:hard

GYPSUM

soft:X:_:_:_:_:hard

Table 2.5. The data matrix.

	T	D	Q	G
Friedrich Moh	4	5	4	1

Table 2.6. The matrix of magnitudes.

	T	D	Q	G	Σ
T	0	0	0	1	1
D	1	0	1	1	3
Q	0	0	0	1	1
G	0	0	0	0	0

the upper right corner. As will be observed, both data matrices of Frame 2.1 are identical to an ideal Guttman-type data matrix, i.e. scales derived from their marginals are ordered in one dimension.

In the next experiment, which will be referenced as the rating scale experiment (Frame 2.2), a prior experience of observation of change is assumed. The data matrix of this experiment is recorded in Table 2.5. The meaning of the elements of this data matrix could be explicated as "hard", "very hard", "hard", and "soft". This meaning is encoded into the data matrix (Table 2.5) as a quantity. Let us now reconstruct the previous hypothetical experience this particular subject had with this sample of minerals. If we adopt a convention that magnitude "greater than" will be recorded as "1" and magnitude "less than" as "0" and compare all elements of the data matrix in Table 2.5, the resulting matrix of magnitudes will be as shown in Table 2.6. Compare the "hardness" scales from both experiments: the N=1 pair comparison experiment yielded the scale of magnitudes [0,1,2,3;G,Q,T,D], which can be directly interpreted as the increasing number of changes observed. The scale of quantities from the rating scale experiment [1,4,4,5;G,T,Q,D] is clearly dependent on the size of the rating interval, and the reconstructed scale of magnitudes [0,1,1,3;G,Q,T,D] is influenced by this 1 to 5 rating interval.

FRAME 2.3. THE N=1 DICHOTOMOUS ITEMS EXPERIMENT.

Below is a series of statements pertaining to the hardness of minerals. Please indicate if you agree or disagree with these statements.

Is the topaz hard?	YES	NO
Is the diamond hard?	YES	NO
Is the quartz hard?	YES	NO
Is the gypsum hard?	YES	NO

Table 2.7. The data matrix.

	T	D	Q	G
Friedrich Moh	1	1	1	0

Table 2.8. The matrix of magnitudes.

	T	D	Q	G	Σ
T	0	0	0	1	1
D	0	0	0	1	1
Q	0	0	0	1	1
G	0	0	0	0	0

FRAME 2.4. The experiment illustrating the effect of increased variance on a matrix of magnitudes.

Below you will find a series of statements pertaining so the hardness of minerals. Please indicate if you agree or disagree with these statements.

1) Is the topaz hard?	YES	NO
2) Is the diamond hard?	YES	NO
3) Is the fluoride hard?	YES	NO
4) Is the gypsum hard?	YES	NO

Table 2.9. The data matrix.

	T	D	F	G
Friedrich Moh	1	1	0	0

Table 2.10. The matrix of magnitudes.

	T	D	F	G	Σ
T	0	0	1	1	2
D	0	0	1	1	2
F	0	0	0	0	0
G	0	0	0	0	0

The third type of administration of our "hardness" test will now be discussed. (Frame 2.3.) The typical data matrix would probably look like the one in Table 2.7 and our reconstruction of previous comparisons will yield the matrix of magnitudes as shown in Table 2.8. The resulting scale of quantities [1,1,1,0;T,D,Q,G] is the least similar to the scale constructed by the pair comparison method [0,1,2,3;G,Q,T,D]. It need not be so. The improvement of sample composition, e.g. exchanging quartz for fluorite (F), could change the data matrix of Table 2.7 into a data matrix reported in Table 2.9 with its corresponding matrix of magnitudes shown in Table 2.10 of Frame 2.4.

Endorsement frequency of monotone items is "a joint function of: (a) the perceived boundary established by the item on the attribute continuum, and (b) the individual's perceived position on this continuum." (Goldberg, 1963, p. 473). Previous "improvement" of the sample composition extended "boundaries" of our test, which was reflected in increased variance (information) content of the matrix from Table 2.10. By extending our sample as in Table 2.11, we have included individuals whose attitudes (and therefore their positions on the attribute measured) differ. This is reflected in an increase of information content of the reconstructed matrix of hypothetic pair comparisons as seen in Table 2.12 (Frame 2.5).

FRAME 2.5 THE N=3 DICHOTOMOUS ITEMS EXPERIMENT

Below you will find a series of statements pertaining to the hardness of minerals. Please indicate if you agree or disagree with these statements:

1)	Is the topaz hard?	YES	NO
2)	Is the diamond hard?	YES	NO
3)	Is the fluoride hard?	YES	NO
4)	Is the gypsum hard?	YES	NO

Table 2.11 The data matrix.

	T	D	F	G
Subject E	1	1	1	0
Subject M	1	1	0	0
Subject A	0	1	0	0

Table 2.12 The matrix of magnitudes.

	T	D	F	G	Σ
T	0	0	1	2	3
D	1	0	2	3	6
F	0	0	0	1	1
G	0	0	0	0	0

As compared with the matrix of magnitudes based on direct observations of changes by a single observer in the N=1 pair comparisons experiment, the matrix of magnitudes in the N=3 dichotomous items experiment contains magnitudes due to joint (by two or more subjects) observation of change. Also, in comparing the scales derived from the N=1 pair comparisons experiment and the N=3 dichotomous items experiment, the former scale [3,2,1,0;D,T,Q,G] increases in steps proportional to the distance (in magnitude units) between observed minerals. The units of the latter magnitude scale [6,3,1,0;D,T,F,G] are products of joint variation between observations and observers. If we are interested in scaling observations, the variance between observers becomes redundant. It should be removed from the final scale and used only for purposes of statistical estimation.

Generation of the difference relation.

Imagine the smooth (D) and scratched, i.e. not smooth ($\overline{D}$) surface of a mineral D. If no operation on these surfaces is taking place, these two types of surfaces are mutually exclusive. The operation of "scratching" can be described as taking place in a series of time intervals t_1, t_2, ... t_i, ... t_n. Intersection I of smooth and scratched surfaces can be written as

$$\lim_{e \to 0} I_e = D^1_{t_{i-1}} \cdot \bar{D}_{t_i}; \text{ where } e = t_{i-1} - t_i \tag{2.1}$$

i.e. as a limit of differences between two temporal intervals of the scratching operation, when the change of the surface actually takes place and the surface is both "smooth" and "not smooth".

In matrix notation, this operation can be written as a matrix of magnitudes M_{qxq} (Table 2.12), generated by products of its component vectors of the data matrix D_{rxq} (Table 2.11), as shown in equations 2.2, 2.3, and 2.4.

$$M_{qxq} = {}^{1}C_{qxq} + {}^{2}C_{qxq} + \ldots + {}^{i}C_{qxq} + \ldots + {}^{r}C_{qxq} \tag{2.2}$$

where

$${}^{i}C_{qxq} = {}^{i}D^{1}_{1xq} \cdot {}^{i}\bar{D}_{1xq} \tag{2.3}$$

and

$${}^{i}\bar{D}_{1xq} = {}^{i}D_{1xq} + U_{1xq} \text{ (mod 2)} \tag{2.4}$$

Component matrices ${}^{i}C_{qxq}$ of the matrix of magnitudes M_{qxq} are composed out of major products of its constituent vectors ${}^{i}D_{1xq}$ and ${}^{i}\bar{D}_{1xq}$. Elements of the vector ${}^{i}D_{1xq}$ of a subject i can be imagined as standing for hypothetical observations of a scratched surface.

Reflected vector ${}^{i}D_{1xq}$ can be written as ${}^{i}\bar{D}_{1xq}$. This reflection is described in equation 2.4 as an addition of a unit vector U_{1xq} in modulo two. Addition of this modulo two unit vector changes elements of the vector ${}^{i}D_{1xq}$ from one to zero and vice versa. Reflected vector ${}^{i}\bar{D}_{1xq}$ can be imagined at our conceptual level of hypothetical observations as describing observations of a smooth surface.

Reconstruction of components ${}^{i}C_{qxq}$ of the matrix of magnitudes M_{qxq} for subjects of the dichotomous items experiment (Tables 2.11 and 2.12) is done in Table 2.13 a,b,c.

Application of principles developed in equations 2.2, 2.3, and 2.4 to entities of a data matrix D_{rxq} leads to the following complementary system of equations:

$$M_{rxr} = \sum_{j=1}^{q} {}^{j}C_{rxr} \qquad (2.5)$$

where

$${}^{j}C_{rxr} = {}^{j}D_{rx1} \cdot {}^{j}\bar{D}^{1}_{rx1} \qquad (2.6)$$

Generation of matrices of magnitudes from their ipsative and reflected component vectors as described by equations 2.2 through 2.5 is based on a primitive concept of binary variation as a number of one-zero changes. This operation can be also conceptualized as a product of a matrix and its negation, which

Table 2.13 a, b, c. Matrices of magnitudes as major products of their constituent vectors. Responses of subjects E, M, and A to dichotomously scored monotone items T, D, F, and G are from Table 2.11.

SUBJECT E:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{(a)}$$

SUBJECT M:

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{(b)}$$

SUBJECT A:

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{(c)}$$

in a static sense should logically result in an empty set. This contradiction was resolved by postulating the intersection of these otherwise mutually exclusive sets as taking place in a temporary infinitesimal interval, when both sets are at the same time both A and non-A; i.e. when the change transpires.

Difference relation, information, and variance.

The case in which the variance is computed for a variable that takes on only two possible values, as in the case of a pass-fail test item, the variance can be computed as

$$\sigma_x^2 = mn \tag{2.7}$$

where m equals the proportion of people which pass the item x and n to the proportion of persons which fail the item x. In this case n = 1-m. (cf. Nunnally, 1967, p. 106). If we define proportion of "ones" (m) of a variable vector of a 1xq order as a/q and proportion of "zeros" (n) of the same vector as b/q, we can write equation 2.7 as

$$\sigma_x^2 = \frac{ab}{q^2} \tag{2.8}$$

As may be observed, order of square matrix ${}^{1}C_{qxq}$ from equation 2.3 is equivalent to the denominator of the right term of the equation 2.8. Magnitude of this matrix is at the same time determined by the product ab (the numerator of equation 2.8).

Consider, as an example, vector ${}^{1}D_{1xq}$ of the Table 2.13a; proportion of "ones," m = 3/4, proportion of "zeroes," n = 1/4 and variance σ_d^2 (equation 2.7) mn = 3/16. Magnitude of the vector ${}^{1}D_{1xq}$ is defined by elements of the ${}^{1}C_{qxq}$ matrix and at the same time defined as 3 by the numerator of equation 2.8. Order of this matrix, determined by the denominator of equation 2.8, equals 4 times 4.

Consider again the matrix of magnitudes from Table 2.13a. An alternative way to obtain this matrix is to view vector ${}^{1}D_{1xq}$ as a marginal referent for both dimensions of the square matrix ${}^{1}C_{qxq}$, i.e. to define margins of the ${}^{1}C_{qxq}$ by ${}^{1}D_{1xq}^{d1}$ and ${}^{1}D_{1xq}^{d_2}$, where ${}^{1}D_{1xq}^{d_1} = {}^{1}D_{1xq}^{d_2}$. If we define elements ${}^{C}i_{d_1}$, i_{d_2} of the matrix ${}^{1}C_{qxq}$ as equal to 1 if ${}^{1}D_{1xq}^{d_1} > {}^{1}D_{1xq}^{d_2}$ and as equal to 0 otherwise (Table 2.14), the resultant matrix will be identical to the matrix obtained by operations described in equations 2.3 and 2.8.

The information content of the ${}^{1}C_{qxq}$ matrix can be computed as

$$H = \sum_{e=1}^{m} \sum_{a=1}^{k} C_{ea} \tag{2.9}$$

where H stands for the amount of information in bits. For our example (Table 2.14) the information content H of the ${}^{1}C_{qxq}$ matrix equals 3 bits.

Table 2.14. Schematization of the concept of binary variance as a number of one-zero changes "within" a variable.

$$
\begin{array}{cc|cccc}
 & & \multicolumn{4}{c}{{}^{1}D^{d_2}_{1xg}} \\
 & & 1 & 1 & 1 & 0 \\ \hline
{}^{1}D^{d_1}_{1xg} & 1 & 0 & 0 & 0 & 1 \\
 & 1 & 0 & 0 & 0 & 1 \\
 & 1 & 0 & 0 & 0 & 1 \\
 & 0 & 0 & 0 & 0 & 0
\end{array}
\qquad
C_{i_{d_1}, i_{d_2}}
\begin{array}{l}
= 1 \text{ if } {}^{i}D^{d_1}_{1xg} > D^{d_2}_{1xg} \\[2em]
= 0 \text{ if } {}^{i}D^{d_1}_{1xg} \leq D^{d_2}_{1xg}
\end{array}
$$

Chapter 3

RELATIONS BETWEEN COGNITIVE, LOGICAL, AND GEOMETRIC STRUCTURES

Cognitive and logical structures.

The notion that a study of cognitive systems is possible if formal systems reflecting natural systems are scrutinized, was expressed by George Boole in his "Investigation of the Laws of Thought." He observes that "whether we regard signs as the representatives of things and of their relations, or as the representations of the conceptions and operations of the human intellect, in studying the laws of signs, we are in effect studying the manifest laws of reasoning." (Boole, 1854, p. 26). Boole's proposal can be reversed and an attempt made to use known general laws of cognitive processes for analysis of heuristic formal systems.

Assume that the forms, provided by the earth, influenced formation of our cognitive systems. Suppose then, that language is used as an instrument for manipulation of signs and their mutual interrelationships. The body of knowledge, describing operations on signs congruent with "normal" functioning of our cognitive systems (adequately reflecting the outside reality) is usually treated as a subject of formal logic.

Corollaries to this argument connecting "normal" type of operations on signs with a correct reflection of objective reality are findings pertaining to "abnormal" types of cognitive functioning. Thus e.g. Sullivan's (1953) prototaxic and parataxic types of cognitive processes "have no necessary connections among themselves and ... are not logically related." (Hall and Lindzey, 1957, p. 140).

Logic is concerned with forms; forms emerge from abstracted properties of contents of experience. Alfred North Whitehead comments on this process: "we ascribe the origin of these sensations to relations between the things which form the external world [and] we want to describe the connections between these external things in some way which does not depend on any particular sensations, nor even on all the sensations of any particular person [...] thus it comes about that, step by step, and not realizing the full meaning of the process, mankind has been led to search for a mathematical description of the properties of the universe, because in this way only can a general idea of the course of events be formed, freed from references to particular persons or to particular types of sensation." (Whitehead, 1911, pp. 4-5). In this process, concrete things are replaced by formalized elements of variable meanings.

Any symbolic structure can be thought of as a structure of propositions if it contains a symbol understood to represent a relation. A proposition asserts that a certain relation holds among certain elements. The dominance matrix described earlier can be thought of as a symbolic structure describing relations between propositions contained by elements of its generic data matrix: a sentence of ordinary language can be considered a proposition if it describes some relation. Every proposition is either true or false. This property is the only connecting link between separate propositions as considered from a formal logic point of view. Thus two sentences are considered algebraically not equal if they are not logically equivalent. Two sentences are not equal if one has the truth value "true" and the truth value of the other is "false", or vice versa. Thus, in a formal sense, a matrix of magnitudes derived from a binary data matrix, may be considered a description of relations between the data matrix elements.

The sentence "I stay in the background" (Pemberton, 1952) contains a variable "I"; it can be regarded as a function of a particular person i from a set of persons 1,2, ...,i, ..., n encountering this sentence. Because the function f(i) becomes a definite sentence (a proposition) for each value of i, it is called a propositional function. The value of each function f(i) can be denoted $f(i_n)$ so that, for example, $f(i_4)$ means that person number four stays in the background, which can have the truth value of one ("true") or zero ("false").

A calculus of bivalued truth values consists of sixteen propositional functions, summarized in Table 3.1. These functions can be classified as reflexive and aliorelative classes. A relation is reflexive provided that its truth value is a function of a singular term. A reflexive class of truth functions includes the relations of tautology $A V \bar{A}$, contradiction A&A and relations such as A&A, $\overline{A \& A}$, B&B and $\overline{B \& B}$.

The truth values of aliorelative functions depend on the combination of truth values of their nonsingular terms. A relation is aliorelative if no term has this relation to itself. Given any serial relation (say R), no term must precede itself. (i.e. xRx is not valid.)

Aliorelative functions can be further subdivided into order-dependent and order-independent classes. The order-independent class of aliorelative functions consists of functions for which the order of their generic arguments does not make a difference, i.e. relations which are both aliorelative and symmetric, as the relations of conjunction, disjunction, equivalence and their converses.

The order-dependent-aliorelative functions are asymmetric. This class of propositional functions contains the family of implicative functions, i.e. implication, negative implication, and their converses. As depicted in Table 3.1, if we delete all reflexive functions, the family of implicative functions

Table 3.1. Propositional functions of a bivalued logical calculus. The truth value "1" denotes "true" and the truth value "0" denotes "false". Column "a" contains all possible patterns of (a plenum) of truth values for both arguments A and B. The names of propositional functions are:

REFLECTIVE CLASS

COL.
b) Tautology
e) Ipsative A
g) Ipsative B
l) Reflected B
n) Reflected A
g) Contradition

ALIORELATIVE ORDER-INDEPENDENT CLASS

COL.
c) Disjunction
h) Equivalence
i) Conjunction
j) Sheffer Stroke
k) Nonequivalence
p) Joint Denial

ALIORELATIVE ORDER-DEPENDENT CLASS

COL.
d) Converse of Implication
f) Implication
m) Negative Implication
o) Converse of Negative Implication

A B	$A \vee \bar{A}$ $B \vee \bar{B}$	$A \vee B$	$A \leftarrow B$	$A \,\&\, A$	$A \rightarrow B$	$B \,\&\, B$	$A \equiv B$	$A \,\&\, B$	$A \mid B$	$A \not\equiv B$	$\overline{B \,\&\, B}$	$A \rightarrow B$	$\overline{A \,\&\, A}$	$A \leftarrow B$	$A \downarrow B$	$A \,\&\, \bar{A}$ $B \,\&\, \bar{B}$
1 1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
1 0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0
0 1	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0
0 0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q

will be the only one showing one-zero changes reflecting the one-zero changes in their arguments, a property called in other contexts information or variance.

Dimensionality and order.

Russell (1903) arrived at the definition of the order relation as the relation having asymmetric, transitive, and connected properties. Magnitude is a primitive term with respect to asymmetry, which is the development of the concept of magnitude at the functional level. The formal definition of asymmetricity is accomplished by the introduction of the concept of the square of a relation.

The square of a relation is a relation between two terms that holds between two terms and their intermediate term. Asymmetric relation then can be defined as a relation with an aliorelative square. In a similar vein, transitive relation can be described as a relation which contains its square.

The property of connectedness can be described as the possibility of shared-boundaries-arrangement. The formal definition given by Hempel is "if a does not coincide with b, then a precedes b or b precedes a" (Hempel, 1952, p. 59). The classical example of a nonconnected set of numbers is a set of complex numbers. Every complex number is jointly defined by its real and imaginary parts. It is impossible to decide which part should be used for generation of an ordered progression in exclusion of the other part. For this reason it is impossible to order a set of complex numbers in one dimension.

In Euclidian geometry, a unidimensional space is imagined as a line, defined as a "breadthless length." (Euclid, T. L. Heath's edition, 1926, p. 153). A logical definition of a line, avoiding equivocality of "breadthless" is possible; this definition is based on properties of an order relation. There are several ways to generate an order relation (Russell, 1903, pp. 199-233, 371-380). Imagine a set of elements where every element e_i is either greater or smaller (as recorded with respect to a singular property) than another element e_{i+n}. Mutual relations between a set of such elements will then be asymmetrical and transitive. Any pair of such elements (e_i; e_{i+n}) can be represented as a segment. Addition or subtraction of these segments, representing singular property, will result in a configuration of points having one degree of freedom, i.e. in a line in one dimension. A series of two dimensions can be generated if it is possible to divide the above set of elements into two subsets i and j; where every pair of each subset (e_i; e_{i+n}) and (e_j; e_{j+n}) can be arranged into asymmetrical and transitive series. The formal definition for generation of n dimensional series is given by Russell (1903, p. 375) as follows: "Let there be some series u_1, whose terms are all themselves serial relations. If x_1 be any term of u_1, and x_2 any term of the field of x_1, let x_2 be again a serial relation, and so on. Proceeding to x_3, x_4, etc. let x_{n-1}, however obtained, be always a relation generating a simple series. Then all the terms x_n belonging to the field of any serial relation x_{n-1}, form an n-dimensional series."

A spatial model of these n-dimensional series can be therefore constructed from the potentiality of the data matrix for order. The set of dimensions of a given data matrix is equivalent to the set of ways data can be ordered. In this sense, "dimensions [are] a development of order." (Russell, 1919, p. 29).

Logical and geometric structures.

As seen in the previous section, order can be defined as a condition of logical arrangement among elements of a data matrix. Consider the possibility of an order structure, reflecting a simple logical (and ultimately cognitive) structure. One of the basic logical structures was described by Aristotle as a syllogism. A traditional syllogism can be written as a conjunction of implication functions, i.e. as $(A \rightarrow B)$ & $(B \rightarrow C)$. This implicative chaining is at the core of both syllogistic reasoning and the generation of a straight line-dimension

(Russell, 1903, pp. 382-384). The process of exploring the logical implications is thus a form of research and discovery, i.e. a process of valid inferences based on implicative chaining (Cohen and Nagel, 1934).

Table 3.2 illustrates the generation of a dimension by one type of cognitive structuring of reality, formalized by the rules of syllogistic reasoning. Table 3.2 was constructed by considering all possible arrangements of truth values for arguments A, B, and C in step 1, recording the truth values of implications for the arguments A → B in step 2 and for arguments B → C in step 3. Truth values generated in steps 2 and 3 were joined by the conjunctive function in step 4. As designated earlier, implication (→) is false ("0") only if the conclusion is false ("0") and premise true ("1"). This corresponds at the data level (step 1) to the response pattern (1,0). Conjunction (&) is true only in the case of a (1,1) response pattern.

It is feasible to imagine the set of all possible arguments of truth values for arguments A, B, C as a set of all possible response patterns (a plenum), which can be divided into unidimensional subsets of response patterns. The subset of response patterns compatible with the logic function (A → B) & (B → C) was formed in step 4, where compatible response patterns (i.e. patterns leaving their corresponding values in

Table 3.2 Generation of a dimension by a logic function (A → B) & (B → C). A three dimensional plenum of three variables was constructed in step 1. Its one dimension, recorded in step 5, was extracted in steps 2, 3, and 4.

POSSIBLE RESPONSE PATTERNS	LOGICAL STRUCTURE			COMPATIBLE RESPONSE PATTERNS
A B C	(A → B)	&	(B → C)	A B C
1 1 1	1	1	1	1 1 1
1 1 0	1	0	0	———
1 0 1	0	0	1	———
1 0 0	0	0	1	———
0 1 1	1	1	1	0 1 1
0 1 0	1	0	0	———
0 0 1	1	1	1	0 0 1
0 0 0	1	1	1	0 0 0
Step 1	Step 2	Step 4	Step 3	Step 5

step 4 equal to "1") were determined and recorded in step 5.

The set of response patterns, generated by the function (A → B) & (B → C) is equivalent to a Guttman-type data matrix, i.e. to such a matrix whose marginal sums would conform with an ideal Guttman scale (Guttman, 1944). Unidimensionality of this type of data matrix was proven by Torgerson (1958, p. 312).

Since upper (1,1,1 ...) and lower (0,0,0 ...) boundaries of Guttman scales transmit no information, a simpler model of syllogistic cognitive structure can be constructed, taking into account only one-zero and zero-one binary tuples; i.e. this type of analysis is based on frequencies of the negative implication function and its converse. The negative implication is true only in the case of the (1,0) ordered tuple and false in the case of tuples (i,1), (0,1), and (0,0). The converse of negative implication is true only in the case of the (0,1) ordered tupe and false in the case of tuples (1,1), (1,0), and (0,0). This model is computationally simpler and results in series of Guttman scales equivalent to the model schematized in Table 3.2.

Conceptual relations between order and factor analyses.

The idea to consider the logic functions of propositional calculus as variance-generative is derived from Boole (1854) and can be ultimately traced to Aristotle: "...for truth and falsity imply combination and separation ..." (Aristotle, W. D. Ross edition, 1927, pp. 7-8). It is possible to observe the affinity of this notion of "combination and separation" as necessary conditions for the derivation of logical truth or falsity with those of information theory (Shannon and Weaver, 1949), defining information as implying change.

The postulation of correspondence between cognitive processes and logical functions is credited to Aristotle, who considered the rules of syllogism and its conversion into (logical) figures to define elementary processes of all reasoning (Jaeger, 1923). Boole's (1854) analysis of relations between premises of syllogism was motivated by his conviction that this type of activity will ultimately lead to a general formula for solving abstract problems.

Leibnitz, searching for a universal method to "enhance the capabilities of the mind" (Leibnitz, P. P. Wiener's edition, 1951, p. 23) proposes to reduce this problem to the "solution of equations whose roots must be extracted analytically by means of calculation, or geometrically by means of the intersection of loci" (Leibnitz, P. P. Weiner's edition, 1951, p. 5). Later he remarks that "Mind, not contented with agreement, conceives the application of relations containing a certain order" (Leibnitz, P. P. Wierner's edition, 1951, p. 253).

Note that the former proposed solution to the problem of analytic discovery is in essence the description of the process of extraction of eigenvalues in factor analysis, where the roots of a characteristic equation define the initial axes of the test space.

As a matter of fact, these historical antecedents and analogies were secondary with respect to our initial effort to translate the characteristic equation of factor analysis into formal logic functions. The initial search centered around the possibilities of supplanting the roots of a characteristic equation, determining the initial position of a dimension (lines) in a multidimensional space by means of linear scales derived from Guttman-type data matrices. This in turn led to our reconsideration of the Spearman tetrad-difference criterion and the Thurstone matrix algebra formulation of a tetrad as a second order minor expansion.

Thurstone describes his matrix algebra extension of the tetrad criterion in the preface to the "Multiple Factor Analysis": "In 1931 I decided to investigate the relation between multiple factor analysis and Spearman's tetrad differences. When I wrote the tetrad equation to begin this inquiry, I discovered that the tetrad was merely the expansion of a second-order minor, and then the relation was obvious. One might speculate as to whether multiple-factor analysis would have developed earlier if this interpretation had been stated earlier. If the second-order minors must vanish in order to establish a single common factor, then must the third-order minors vanish in order to establish two common factors, and so on? To have put the matter in this way would have led to the matrix formulation of the problem much earlier, as well as to the immediate development of multiple-factor analysis. Instead of dealing with the proportional columns and rows of a hierarchy and the vanishing tetrads, we now deal with the same relations in terms of the properties of unit rank, namely, proportional columns and rows and vanishing second-order minors." (Thurstone, 1947, p. vi.)

Spearman describes his criterion of "tetrad differences" in the following words: "The start of the whole inquiry was a curious observation made in the correlations calculated between the measurements of different abilities. These correlations were noticed to tend towards a peculiar arrangement, which could be expressed in a definite mathematical formula

$$r_{ap} \times r_{bq} - r_{aq} \times r_{bp} = 0.$$

This formula has been termed the tetrad equation and the value constituting the left side of it is the tetrad difference. An illustration may be afforded by the following imaginary correlations between mental tests. For instance, let us try the effect of making:

Table 3.3 Hypothetical correlations between four mental tests. (Reproduced from Spearman, 1932, p. 74).

	Opposites	Completion	Memory	Discrimination	Cancelation
Opposites	—	.80	.60	.30	.30
Completion	.80	—	.48	.24	.24
Memory	.60	.48	—	.18	.18
Discrimination	.30	.24	.18	—	.09
Cancellation	.30	.24	.18	.09	—

a denote Opposites, b denote Discrimination, p denote Completion, and q denote Cancellation.

From the table of correlations above, we see that r_{ap} will mean the correlation between opposites and completion, which is .80. Obtaining in a similar fashion the other three correlations needed, the whole tetrad criterion becomes

$$.80 \times .09 - .30 \times .24 = 0$$

which is obviously correct. And so will be found any other application whatever of the tetrad equation to this table." (Spearman, 1932, pp. 73-74).

Trying to link Spearman's criterion of tetrad differences, Thurstone's notion of the rank of the matrix as its dimensionality indicator, and the unidimensional Guttman scale, a discrepancy arose. The Guttman scale was considered unidimensional, but its generic data matrix was not. As a matter of fact, Guttman (1954) based his theory of scalable attitudes on the principal component of his data matrix and maintained that "a sample of ten dichotomous items from a perfect scale can yield at most eleven distinct ranks. In the universe of items, there may be an infinite number of distinct ranks." (Guttman, 1950, p. 296).

Even though Torgerson called attention to the fact that "the matrix of interitem tetrachorics [in a perfect Guttman scale] is a matrix of positive ones" (Torgerson, 1958, p. 312), he considered this finding to be important mainly with respect to properties of its constituent types of correlation coefficients: "such cannot be said, however, of the matrix of point correlations between items, [in which the] rank will be equal to its order." These findings were recognized by Lord and Novick: "the number of common factors in a correlation matrix depends on the type of correlation coefficient used. It also depends on how the item scores are transformed before the correlations are computed." (Lord and Novick, 1968, p. 382). The same authors recognize the priority of the "dimensionality of the complete latent space" over the number of common factors as derived by factor analysis, but at the same time concede that this problem "has not been completely solved." (Lord and Novick, 1968, p. 382).

During the attempts to increase the amount of information conveyed by the ideal Guttman scale we tried various monotonic transformations of its elements. Logical intersections of the marginals seemed to be promising because of their compatibility with the overall sequential arrangement of the Guttman type data matrix. (Table 3.4 and 3.5) A routine check of the tetrad of this transformed matrix (Table 3.5) showed it to be zero:

Table 3.4 Guttman-type data matrix and its marginals.

A	B	C	
1	1	1	$\frac{3}{3}$
1	1	0	$\frac{2}{3}$
1	0	0	$\frac{1}{3}$
0	0	0	$\frac{0}{0}$
$\frac{3}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	

Table 3.5 A matrix of products of Guttman-type data matrix marginals.

$\frac{9}{12}$	$\frac{6}{12}$	$\frac{3}{12}$
$\frac{6}{12}$	$\frac{4}{12}$	$\frac{1}{12}$
$\frac{3}{12}$	$\frac{2}{12}$	$\frac{1}{12}$
$\frac{0}{12}$	$\frac{0}{12}$	$\frac{0}{12}$

$$\frac{\frac{9}{12}}{\frac{6}{12}} - \frac{\frac{6}{12}}{\frac{6}{12}} = 0$$

$$\frac{9 \times 6}{12 \times 12} - \frac{6 \times 6}{12 \times 12} = 0$$

$$\frac{36}{144} - \frac{36}{144} = 0$$

$$0 = 0$$

This finding, connecting the transformed Guttman-type data matrix and Spearman's tetrads (and thus Thurstone's ranks) fostered our belief that the reconceptualization of the characteristic equation in terms of formal logic structures is possible. A similar belief was expressed by Thurstone himself, who considered the reconceptualization of the factor analytic model similar to order analysis, as one of the possibilities for the future development of the factor analytic method: "Analysis of successive differences [should] lead to the factor pattern, which would indicate the number of parameters involved in the variance of each test. That is the main object of factor analysis [...] in teasing out what we have called the underlying order of a new domain" (Thurstone, 1947, p. xiv).

Chapter 4

THE ELEMENTS OF [illegible]ER ANALYSIS

The prerequisite and disconfirmatory relationships.

There are several ways to search for relationships between the data matrix elements. It is possible to analyze every method in terms of its underlying and implicitly assumed logical constants and functions. Within the system of logical constants, various constituent relationships involving logical connectives as e.g. "and," "if and only if," "either," "either... or," may be employed. As seen earlier, the family of implication functions is dimensionality-generative and therefore of interest to behavioral researchers. The logical connectives of the family of the implication functions (Table 3.1., d, f, m, o) can be written as (d) "is a prerequisite to," (f) "implies," (o) "is not a prerequisite to," (m) "does not imply."

Within the family of implicative functions, it is possible to move from one function to another by interchanging or reflecting variables within the systems. Close scrutiny of this family will also reveal that (1,0) and (0,1) arguments of these functions are variance-generative, while the (1,1) and (0,0) tuples are important in the determination of the orientations of variables within the system. If no attempt to optimize a variable's orientation is made, the simplest functions suitable for analysis are the functions of negative implication and its converse. These two functions differ only in their (1,0) and (0,1) tuples. The (1,0) tuple will be called a prerequisite (or confirmatory) response pattern. The (0,1) tuple will be called a disconfirmatory response pattern.

The logic diagram of a test space (the nonmetric model A).

Using only the prerequisite and disconfirmatory response patterns, a logic diagram (tree) of manifest structures can be constructed. Before formalizing this process, let us consider a simple example.

A twelve item rating scale of guilt, constructed *ad hoc*, was administered to fifteen students enrolled in a general psychology course at the University of Minnesota. Instructions were printed in caption on the rating scale: "Imagine that you find yourself in the situations described below. Rate how you would feel if it happened. Be frank." The answer choices to each item were the following:

: very bad : a little bad : not too bad : don't care :

The items were scored in a bivalued manner with "1" being given to either of the first two choices and "0" being given to either of the last two choices. Table 4.1 lists the items presented to the subjects. The data matrix is presented in Table 4.2.

First, write randomly letters A through L on a piece of paper as:

Table 4.1. The guilt scale.

A. Drinking too much

B. Shoplift

C. Cheat on exams

D. Have an extramarital affair

E. Gossip

F. Don't go to church on Sunday

G. Have an homosexual experience

H. Lie to parents

I. Lie on income tax returns

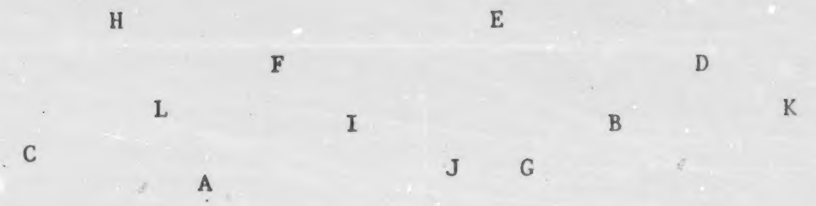
J. Being caught as a Peeping Tom

K. Steal a book from library

L. Steal a book from a friend

Table 4.2 Data matrix for twelve-item scale of guilt for sample of fifteen subjects.

Subject Item	A	B	C	D	E	F	G	H	I	J	K	L	row totals
1	1	1	1	1	0	0	1	1	1	1	1	1	10
2	1	1	0	1	1	0	1	0	1	1	1	1	9
3	1	1	1	1	0	0	1	1	1	1	1	1	10
4	0	1	0	0	0	0	1	0	0	1	0	1	4
5	0	0	1	0	1	0	1	0	0	1	1	1	6
6	1	1	1	1	1	0	1	1	1	1	1	1	11
7	0	1	1	1	1	0	1	0	1	1	1	1	9
8	0	1	0	1	1	0	0	1	0	0	0	1	5
9	0	1	1	0	0	0	1	1	0	0	0	1	5
10	1	0	0	0	0	0	1	0	1	1	0	1	5
11	1	1	0	1	0	0	1	0	0	1	0	1	6
12	0	1	0	1	1	0	0	0	1	1	1	1	7
13	0	1	1	1	1	0	1	1	0	1	0	1	8
14	1	1	0	1	0	0	1	0	1	1	1	1	8
15	0	1	1	1	0	0	1	0	1	1	1	1	8
column totals	7	13	8	11	7	0	13	6	9	13	9	15	



Next, start tree construction with item A and Subject 1 (Table 4.2). For Subject 1, item A will be a prerequisite for items E and F (response patterns "1,0"). For Subject 2, item A will be a prerequisite for items C, F, and H. For Subject 3, item A will be a prerequisite for items E and F. Finally, for Subject 4, who answered item A in a negative sense, items B, G, J, and L will be the prerequisites. Item A is therefore the prerequisite for items E, F, C, H, B, D, K, and I (read across subjects). Items B, G, J, L, C, E, K, D, I, H are prerequisites for item A. The prerequisite pattern "1,0" conflicts with the disconfirmatory patterns "0,1" for all items except item F. We now can draw a line connecting items A and F.

The preceding steps can be supplanted by looking for patterns of zeros across all subjects: (1) start with the first row, containing zero in the scrutinized item's column (2) write down all items containing zero in this row (3) find next row containing zero in the scrutinized item's column (4) read across this row and underline every previously marked item with "1" in this row (5) continue to the last row (6) start a new item, repeat steps one through six until all items are scrutinized (7) draw lines, connecting scrutinized item with items not underlined (8) rearrange items so their connecting lines do not cross. Steps (1) through (6) are written below:

ITEM A $[C,\underline{D},\underline{E},F,\underline{H},I,\underline{K}]$ (A → F)

ITEM B $[\underline{A},D,F,H,\underline{T}]$ (B → D,F,H)

ITEM C $[F,\underline{H}]$ (C → F)

ITEM D $[\underline{A},\underline{C},\underline{E},F,\underline{H},\underline{I},\underline{K}]$ (D → F)

ITEM E $[F]$ (E → F)

ITEM F $[\underline{E}]$ (F → NONE)

ITEM G $[A,C,F,\underline{I},\underline{J},\underline{K}]$ (G → A,C,F)

ITEM H $[\underline{C}, F]$ (H → F)

ITEM I $[\underline{A},\underline{C},\underline{D},\underline{E},F,\underline{H},\underline{K}]$ (I → F)

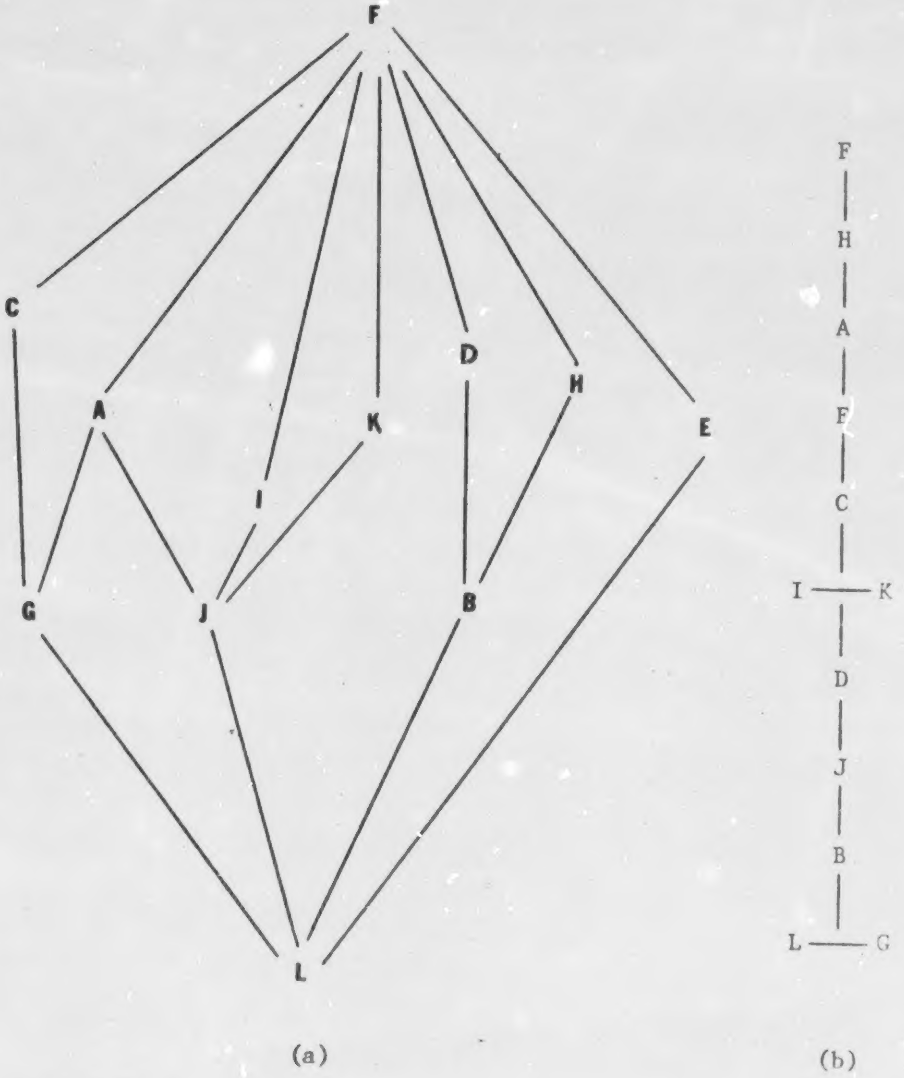
ITEM J $[A,\underline{C},F,\underline{G},I,K]$ (J → A,F,I,K)

ITEM K $[\underline{A},\underline{C},\underline{D},\underline{E},F,\underline{H},\underline{I}]$ (K → F)

ITEM L [ALL] (L → ALL)

The resulting tree is presented in Figure 4.1a. As compared with Guttman's scale constructed for the same data by the "Cornell technique" (Figure 4.1b) Model A tree structure indicates presence of several orders (dimensions) suppressed as error by the Guttman scale.

Figure 4.1. (a) Manifest structure of attributes (data matrix 4.2) isolated by the deterministic Model A of order analysis; (b) The Guttman scale for the same set of items, as constructed by the "Cornell" technique.


F
C
A
I
K
D
H
E
G
J
B
L
(a)
F
H
A
F
C
I
K
D
J
B
L
G
(b)

The logic diagram of a unit test space (Model B).

To systematize the procedure for construction of a data matrix logic diagram, rearrange the data matrix according to its marginals, either in descending (Table 4.3) or ascending order.

As an example of the dominance matrix construction consider items C and H (Table 4.3). Item C is a prerequisite for item H in the case of Subjects 7, 15, and 5, item H is a prerequisite for item C in the case of subject 8 and item H is considered to be equivalent to item C by subjects 6, 1, 3, 2, 13, 14, 12, 11, 10, and 4. Item H is therefore considered a prerequisite for item C by one subject out of 15, item C is a prerequisite for item H for 3 subjects, and item C is considered equivalent to item H by 11 subjects. Thus the probability of a "prerequisite" pattern (10) is 1/15 (.07), and the probability of "disconfirmatory" pattern (01) equals 3/15 (.20). In Table 4.4 probabilities of response patterns (10;01) are listed for all twelve items of the guilt scale.

Asterisks in this matrix's cells signify response patterns confirmatory across all subjects. On the basis of these values, a Model B tree can be constructed (Figure 4.2a) by drawing lines connecting items with asterisks in their corresponding cells. Probability values of elements of this dominance matrix (Table 4.4) can be used as direct indicators of distances between the

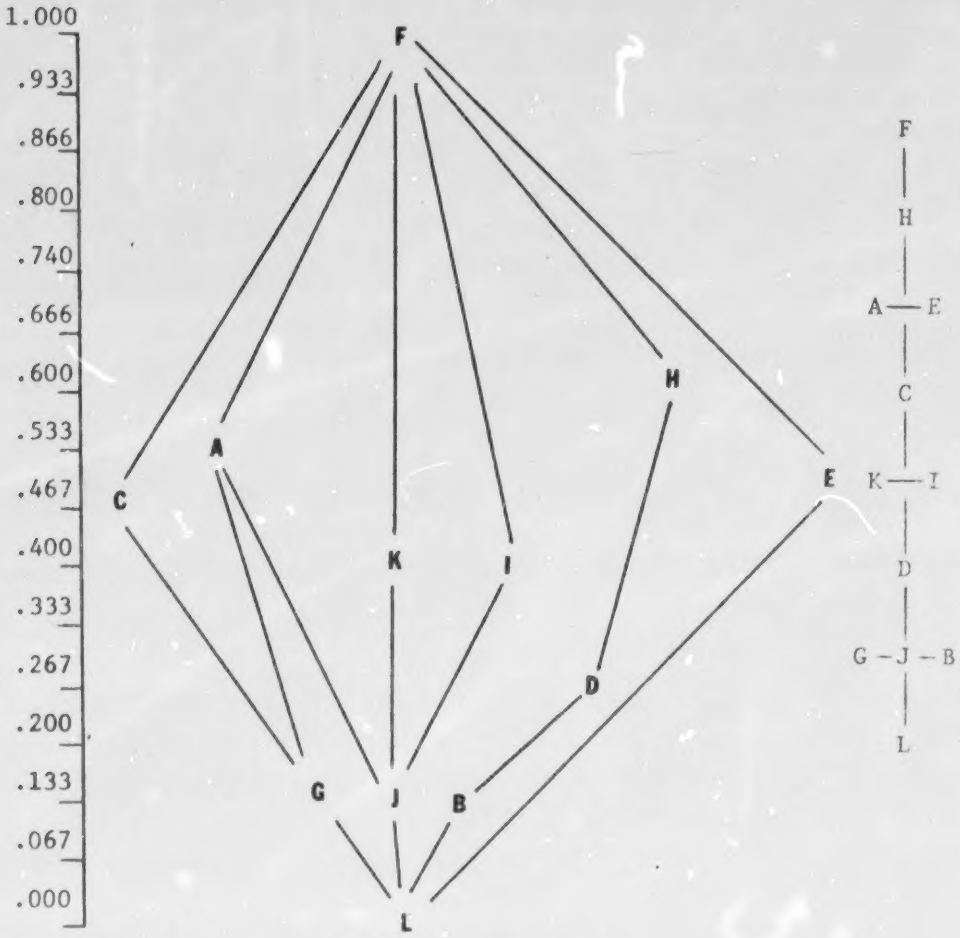
Table 4.3. Reordered data matrix (Table 4.2) for the scale of guilt. Cutting lines, used for construction of Guttman scale, were set as to minimize error (Edwards, 1957, pp. 184-188).

Subject item	L	J	G	B	D	K	I	C	E	A	H	F	row totals
6	1	1	1	1	1	1	1	1	1	1	1	0	11
1	1	1	1	1	1	1	1	1	0	1	1	0	10
3	1	1	1	1	1	1	1	1	0	1	1	0	10
2	1	1	1	1	1	1	1	0	1	1	0	0	9
7	1	1	1	1	1	1	1	1	1	0	0	0	9
13	1	1	1	1	1	0	0	1	1	0	1	0	8
14	1	1	1	1	1	1	1	0	0	1	0	0	8
15	1	1	1	1	1	1	1	1	0	0	0	0	8
12	1	1	0	1	1	1	1	0	1	0	0	0	7
11	1	1	1	1	1	0	0	0	0	1	0	0	6
5	1	1	1	0	0	1	0	1	1	0	0	0	6
8	1	0	0	1	1	0	0	0	1	0	1	0	5
9	1	0	1	1	0	0	0	1	0	0	1	0	5
10	1	1	1	0	0	0	1	0	0	1	0	0	5
4	1	1	1	1	0	0	0	0	0	0	0	0	4
column totals	15	13	13	13	11	9	9	8	7	7	6	0	

Table 4.4. The dominance matrix (Model B) for the data from Table 4.2.

	L	J	G	B	D	K	I	C	E	A	H	F
L	—	.13*	.13*	.13*	.27*	.40*	.40*	.47*	.53*	.53*	.60*	1.00*
J	.00	—	.13	.13	.20	.27*	.27*	.40	.40*	.60	.87*	
G	.00	.07	—	.13	.27	.33	.33	.33*	.53	.40*	.53	.87*
B	.00	.13	.13	—	.13*	.27	.27	.40	.47	.47	.47*	.87*
D	.00	.07	.13	.00	—	.20	.20	.33	.33	.33	.40	.73*
K	.00	.00	.07	.07	.07	—	.07	.20	.27	.27	.40	.60
I	.00	.00	.07	.07	.07	.07	—	.27	.33	.20	.40	.60*
C	.00	.07	.00	.07	.13	.13	.20	—	.27	.27	.20	.53*
E	.00	.07	.13	.07	.07	.13	.20	.20	—	.33	.27	.47*
A	.00	.00	.00	.07	.07	.13	.07	.27	.33	—	.27	.47*
H	.00	.07	.07	.00	.07	.20	.20	.07	.20	.20	—	.40*
F	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	—

Figure 4.2. (a) Manifest structure of attributes from the data matrix 4.2 as isolated by the deterministic model of order analysis B. (b) The Guttman scale for the same set of items as constructed by the Goodenough technique.


1.000
.933
.866
.800
.740
.666
.600
.533
.467
.400
.333
.267
.200
.133
.067
.000
F
H
A
E
C
K
I
D
G
J
B
L
F
H
A — E
C
K — I
D
G – J – B
L

(a) (b)

nodes of Model B tree. Another interpretation of these values is that an item is a prerequisite for another item for *i* subjects out of N.

Dimensionality and reproducibility

The general deterministic procedure for construction of a scale in one dimension is known as scalogram analysis and was described in detail by Guttman (1941, 1942, 1944, 1946, 1947, 1950, 1954). In his formulation, the idea of a perfect scale "affords a rigorous test for the existence of single *meaning* for an area" (Guttman, 1950, p. 88, italics mine) and "requires that each person's responses should be reproducible from his rank alone" (Guttman, 1950, p. 62).

Unfortunately, scalogram analysis does not hold in practice for "any sizable set of items" (Edwards, 1957, p. 181). Attempts to accomodate error by the "Cornell method" (Guttman, 1947) and by the "coefficient of reproducibility" (Guttman, 1940) were criticized as insufficiently operationalized (Festinger, 1947; Loevinger, 1948) and "unwieldy" (White and Saltz, 1957, p. 243). An alternative approach to the problem of errors in univocal scales of measurement will be the topic of this discussion.

Consider a matrix of all possible response patterns to a set of k items (a plenum) as $E_{2^k xk}$ and its corresponding Guttman-type data matrix $G_{k+1 \times k}$. If we disregard the indeterminate boundary patterns (1 1 1 ...) and (0 0 0 ...) of both matrices and assume the unidimensionality of the Guttman scale, the dimensionality of the $E_{(2^k - 2) \times k}$ matrix can be approximated by the ratio of its order to the order of matrix $G_{k - 1 \times k}$, i.e.

$$d_e = \frac{(2^k - 2) \times k}{(k - 1) \times k} = \frac{2^k - 2}{k - 1} \qquad (4.1)$$

Let us designate the response pattern pair (1,1) as η, the "prerequisite" response tuple (1,0) as π, the "disconfirmatory" response pair (0,1) as ε, and the (0,0) tuple as ξ. The average proportion of these response pattern pairs should sum to a unity and can be written as

$$\frac{\sum_{i=1}^{k} \eta_i}{N} + \frac{\sum_{i=1}^{k} \pi_i}{N} + \frac{\sum_{i=1}^{k} \varepsilon_i}{N} + \frac{\sum_{i=1}^{k} \xi_i}{N} = 1 \qquad (4.2)$$

It can be shown that in the $E_{2^k \times k}$ matrix of all possible response patterns, all four coefficients are present in equal proportions, i.e.,

$$\frac{(2^{k-2})\ (k - 1)}{(2^k)\ (k - 1)} = \frac{1}{4} \qquad (4.3)$$

In the Guttman-type data matrix $G_{k+1 \times k}$, the proportion of G_η and G_ξ coefficients is unchanged, and the value of the G_ε coefficient is equal to zero by definition. The value of the

G_{π} coefficient can be calculated as:

$$\frac{\sum_{i=1}^{k} G\pi_{i}}{N} = \frac{k}{2(k+1)} \qquad (4.4)$$

where k is any odd number. When the k approaches infinity,

$$\lim \frac{\sum_{i=1}^{k} G\pi_{i}}{N} = \frac{1}{2};\ k \to \infty \qquad (4.5)$$

For the data matrix (Table 4.2) the number of "1,0" and "0,1" response pairs is recorded in the dominance matrix (Table 4.5). The number of "0,0" and "1,1" response pairs is recorded in the consonance matrix (Table 4.6). Expected proportions of the η, π, ϵ, and ξ, indexes are listed in Table 4.7 and obtained proportions are listed in Table 4.8.

It is possible to test for the significance of difference between obtained and expected means of π, ϵ, and $(\eta+\xi)$ coefficients. The combined index $({}^{\eta}k + {}^{\xi}k)$ can be interpreted as the degree a plenum is proportionally represented in the sample. Coefficient ${}_{\pi k}$ is a measure of the overall tendency of the whole test to approximate the Guttman scale. The elevated coefficient ${}_{\epsilon k}$ indicates the presence of other dimensions in the test space designed by the data matrix.

Table 4.5. The dominance matrix for items from Table 4.2.

	L	I	G	B	D	K	I	C	E	A	H	F
L	—	2	2	2	4	6	6	7	8	8	9	15
J	0	—	2	2	3	4	4	6	6	6	9	13
G	0	1	—	2	4	5	5	5	8	6	8	13
B	0	2	2	—	2	4	4	6	7	7	7	13
D	0	1	2	0	—	3	3	5	5	5	6	11
K	0	0	1	1	1	—	1	3	4	4	6	9
I	0	0	1	1	1	1	—	4	5	3	6	9
C	0	1	0	1	2	2	3	—	4	4	3	8
E	0	1	2	1	1	2	3	3	—	5	4	7
A	0	0	0	1	1	2	1	4	5	—	4	7
H	0	1	1	0	1	3	3	1	3	3	—	6
F	0	0	0	0	0	0	0	0	0	0	0	—

Table 4.6. The consonance matrix for items from Table 4.2.

	L	J	G	B	D	K	I	C	E	A	H	F
L	—	13	13	13	11	9	9	8	7	7	6	0
J	0	—	12	10	10	9	9	7	6	7	4	0
G	0	1	—	11	9	8	8	8	5	7	5	0
B	0	0	0	—	11	8	8	7	6	6	6	0
D	0	1	0	1	—	8	8	6	6	6	5	0
K	0	2	1	1	3	—	8	6	5	5	3	0
I	0	2	1	1	3	5	—	5	4	6	3	0
C	0	1	2	1	2	4	3	—	4	3	5	0
E	0	1	0	1	3	4	3	4	—	2	3	0
A	0	2	2	1	3	4	5	3	3	—	3	0
H	0	0	1	2	3	3	3	6	5	5	—	0
F	0	2	2	2	4	6	6	7	8	8	9	—

Table 4.7. Theoretical proportions of η, π, ε, and ξ coefficients for the $E_{2^k \times k}$ matrix of all possible response patterns to a set of k items, and its corresponding matrix $G_{k+1 \times k}$, representing a Guttman-type data matrix. Coefficients η, π, ε, and ξ, stand for proportions of (1,1), (1,0), (0,1), and (0,0) patterns respectively.

	$E_{2^k \times k}$	$G_{k+1 \times k}$
$p_k(\eta)$	$\frac{1}{4}$	$\frac{1}{4}$
$p_k(\pi)$	$\frac{1}{4}$	$\frac{1}{2}$
$p_k(\varepsilon)$	$\frac{1}{4}$	0
$p_k(\xi)$	$\frac{1}{4}$	$\frac{1}{4}$

Table 4.8. Obtained frequencies and proportions of π, ε, (from Table 4.5) and η, ξ (from Table 4.6) coefficients.

FREQUENCIES

Items	π	ε	η	ξ	$[\eta+\xi]$
L	69	0	96	0	96
J	57	7	87	12	99
G	60	10	86	10	96
B	56	9	86	10	96
D	51	10	80	23	103
K	49	13	69	33	102
I	50	14	68	33	101
C	55	16	59	33	92
E	59	18	48	32	80
A	59	17	52	36	88
H	68	16	43	37	80
F	111	0	0	54	54

PROPORTIONS

Items	π	ε	η	ξ	$[\eta+\xi]$
L	.38	.00	.53	.00	.53
J	.32	.04	.48	.07	.55
G	.33	.06	.48	.06	.53
B	.31	.05	.48	.06	.53
D	.28	.06	.44	.13	.57
K	.27	.07	.38	.18	.57
I	.28	.08	.38	.18	.56
C	.31	.09	.33	.18	.51
E	.33	.10	.27	.18	.44
A	.33	.09	.29	.20	.49
H	.38	.09	.24	.21	.44
F	.62	.00	.00	.30	.30
Sums	4.14	.73	—	—	6.02
Means	.345	.061	—	—	.502

Chapter 5

THE DETERMINISTIC MODEL OF ORDER ANALYSIS

An outline of major computational steps.

The goal of order analysis is the isolation of the manifest and latent structures of the data matrix. In the present form order analysis is restricted to data matrices containing only dichotomous response patterns. These response patterns are conceptualized as atoms in free Boolean algebra (Mendelson, 1970), and the resulting logical net is treated in a way analogous to logical operations performed on truth tables.

In every row and column of the data matrix, the relationship between combinations of pairs of its elements is examined with respect to their truth values generated by the negative implication function and its converse. The number of pairs showing the connected, transitive and asymmetric properties is recorded in the dominance matrix, preserving the direction of asymmetry. If the items were ordered prior to logical analysis according to their decreasing item popularities, the confirmatory response patterns, characterized as one-zero ordered pairs will tend to occur in the superdiagonal part of the dominance matrix. The disconfirmatory response patterns, defined as connected zero-one ordered pairs, will accumulate in the infradiagonal part.

As a next step, elements of the dominance matrix are folded along the main diagonal. A tree structure of the data matrix is formed from those symmetric pairs of the folded dominance matrix which contain zero number of frequencies of the converse of negative implication function. In this limiting case of zero occurrence of disconfirmatory response patterns in adjacent response vectors, branches of the resulting manifest structure of the data matrix are isomorphic with the structure of the Guttman-type data matrices. No provision for error estimation is provided; hence the model is deterministic.

To isolate a latent structure, the data matrix is partitioned into Guttman-type data submatrices. The resulting supermatrix is operated upon in a series of steps. For each submatrix, separate reconstruction of its dominance matrix is performed and resulting Guttman scales are isolated at its marginals. These ordered, unidimensional components of the data matrix are finally concatenated into a composite matrix which is isomorphic with the data matrix latent structure.

Additive solution for the data matrix dimensionality.

The question of a lower bound of dimensionality in the area of deterministic models was first approached by Bennett (1951) and Milholland (1953). Torgerson (1958, p.350) presents a hypothetical set of data which is supposed to fit Coombs' Conjunctive model in two dimensions. Reconstruction of the order of items in separate dimensions is based on Bennett's theorem (Bennett, 1956). To facilitate comparison, order analysis was carried on the same set of data (Table 5.1).

The item arrangement in two dimensions as a result of operations described by Torgerson (1958, p. 351) is reproduced in Figure 5.1. Thus, two-dimensional construction is based on the assumption that data fit into two dimensions as specified by Milholland's formula for a lower bound. It is evident that Coombs' model for this set of data is not completely determined and consequent incompatible patterns are considered errors.

As a first step of order analysis, the data matrix marginals were rearranged in decreasing order. This order can be read from the arrangement of subjects in the dominance matrix S (Table 5.2), and for items in dominance matrix I (Table 5.3).

The supradiagonal of both dominance matrices contain the number of confirmatory responses and the infradiagonal part contains the number of disconfirmatory responses. Using the procedure described previously, the tree was constructed using the data from dominance matrix S (Figure 5.2).

To partition the data matrix into a supermatrix, any branch of the logical tree (Figure 5.2) can be chosen. Let

Table 5.1 Hypothetical data matrix for the conjunctive model. (From Torgerson, 1958, p. 350).

		Items					
		A	B	C	D	E	row totals
Subject types	1	1	0	0	0	0	1
	2	1	1	0	0	0	2
	3	1	1	0	1	0	3
	4	1	1	1	1	0	4
	5	1	1	1	1	1	5
	6	0	1	1	1	1	4
	7	0	0	1	1	1	3
	8	0	0	0	1	1	2
	9	0	0	0	1	0	1
	10	0	1	0	0	0	1
	11	0	1	0	1	0	2
	12	0	1	1	1	0	3
	13	0	0	1	1	0	2
	14	0	0	0	0	0	0
	Column totals	5	8	6	10	4	33

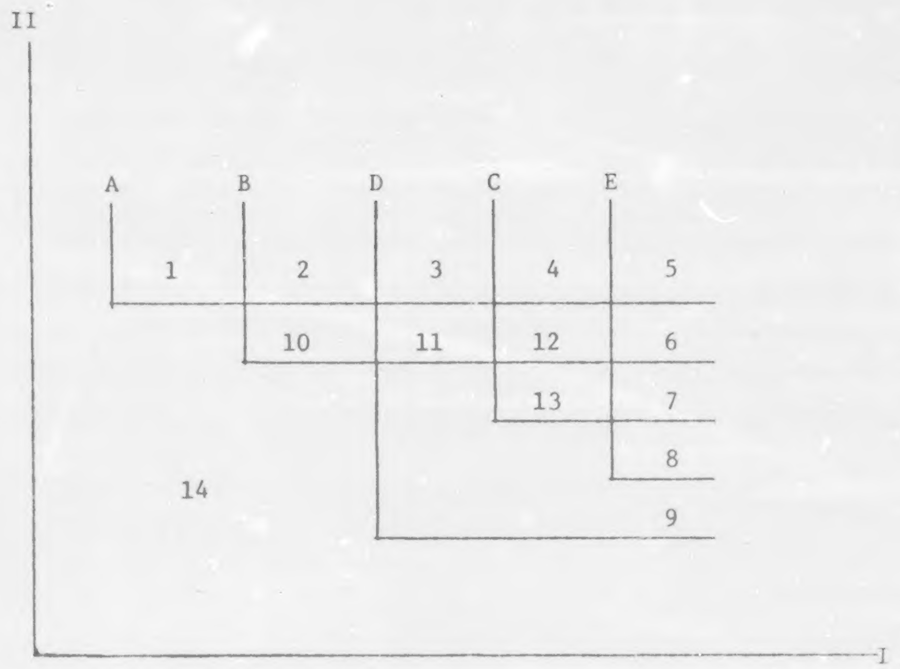


Figure 5.1. Two-dimensional configuration corresponding to the data from Table 5.1. Dimensions are inferred from Milholland's formula for a lower bound of dimensionality and configuration is reconstructed according to Bennett's theorem. Numbered regions in the figure correspond to response patterns of like numbered subject types in Table 5.1. (From Torgerson, 1958, p. 350).

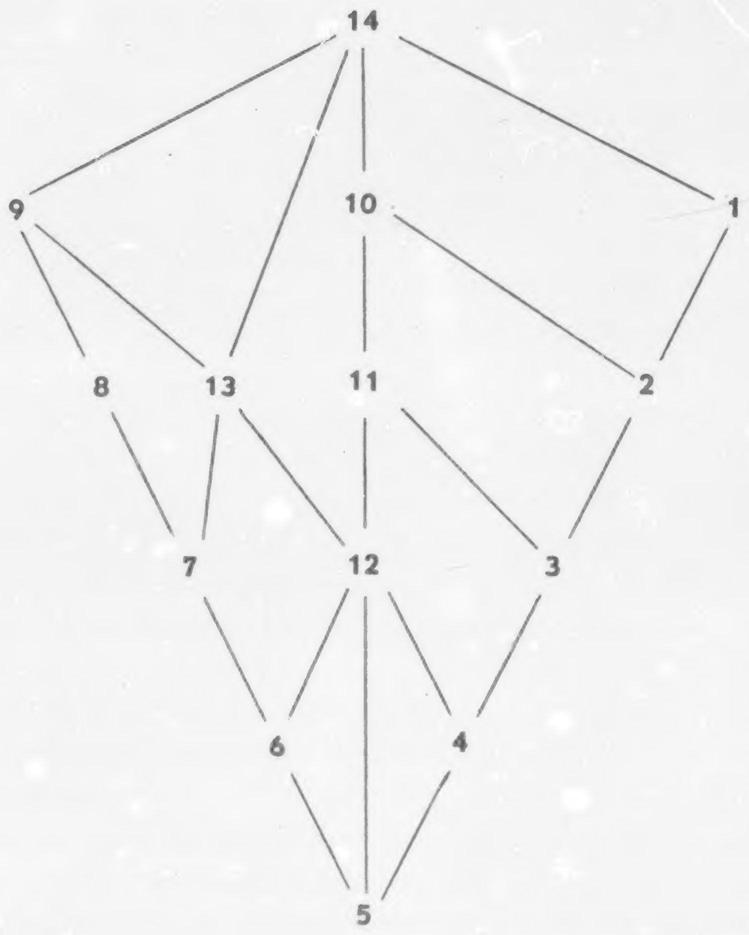
Table 5.2. Dominance matrix S for the data matrix D (Table 5.1).

	Subject types														
Subject types	5	4	6	3	7	12	2	8	11	13	1	9	10	14	Σ
5	—	1	1	2	2	2	3	3	3	3	4	4	4	5	37
4	0	—	1	1	2	1	2	3	2	2	3	3	3	4	27
6	0	1	—	2	1	1	3	2	2	2	4	3	3	4	28
3	0	0	1	—	2	1	1	2	1	2	2	2	2	3	19
7	0	1	0	2	—	1	3	1	2	1	3	2	3	3	22
12	0	0	0	1	1	—	2	2	1	1	3	2	2	3	18
2	0	0	1	0	2	1	—	2	1	2	1	2	1	2	15
8	0	1	0	1	0	1	2	—	1	1	2	1	2	2	14
11	0	0	0	0	1	0	1	1	—	1	2	1	1	2	10
13	0	0	0	1	0	0	2	1	1	—	2	1	2	2	12
1	0	0	1	0	1	1	0	1	1	1	—	1	1	1	9
9	0	0	0	0	0	0	1	0	0	0	1	—	1	1	4
10	0	0	0	0	1	0	0	1	0	1	1	1	—	1	6
14	0	0	0	0	0	0	0	0	0	0	0	0	0	—	0
Σ	0	4	5	10	13	9	20	19	15	17	28	23	25	33	221

Table 5.3. The dominance matrix I for the data matrix D (Table 5.1).

	D	B	C	A	E	Σ
D	-	4	4	7	6	21
B	2	-	4	4	6	16
C	0	2	-	4	3	9
A	21	1	3	-	4	10
E	0	2	1	3	-	6
Σ	4	9	12	18	19	62

Figure 5.2. Manifest structure of the data matrix D (Table 5.1) as isolated by the deterministic model of order analysis.


14
9
10
1
8
13
11
2
7
12
3
6
4
5

us start with the branch composed of subject numbers 5,6,7,8, 9, and 14. Their response patterns are read from the raw score matrix (Table 5.1) and recorded in Table 5.4a. Their item marginals are then computed and items rearranged in descending order as in Table 5.4b. Note that the order of items in the table is identical with D,E,C,B,A order of items in the second dimension of Torgerson's example (Figure 5.1). Table 5.4b can now be imagined as a logical truth table. Negative implication function and the logical functions of equivalence can be computed for every combination of items. The difference values of both functions are recorded in the item dominance matrix for the six subject subset in Table 5.4c. The sum of prerequisite item relationships is listed in the right marginal (column F_I, Table 5.4c). The procedure is repeated for the remaining logical branches of Figure 5.2 and is recorded in Tables 5.4d through 5.41.

At this point of analysis model C will be introduced. As compared with the previously described models A ("nonmetric" and B ("unit space"), model C computations are carried on the matrix elements formed by products of marginals divided by the grand total of the matrix. The data matrix (Table 5.1) is converted into Model C form in Table 5.5. The first factor matrix of Model C is recorded in Table 5.6. Model C data and factor matrices are the monotonic transformations of their Model

Table 5.4. Matrix transformations for factoring the set of hypothetical data from Table 5.1.

	D	B	C	A	E
5	1	1	1	1	1
6	1	1	1	0	1
7	1	0	1	0	1
8	1	0	0	0	1
9	1	0	0	0	0
14	0	0	0	0	0
Σ	5	2	3	1	4

Table 5.4a

	D	E	C	B	A	Σ
	1	1	1	1	1	5
	1	1	1	1	0	4
	1	1	1	0	0	3
	1	1	0	0	0	2
	1	0	0	0	0	1
	0	0	0	0	0	0
Σ	5	4	3	2	1	15

Table 5.4b

	D	E	C	B	A	F_I
D	-	1	2	3	4	10
E		-	1	2	3	6
C			-	1	2	3
B				-	1	1
A					-	0

The
first
factor

Table 5.4c

	D	B	C	A	E
4	1	1	1	1	0
3	1	1	0	1	0
2	0	1	0	1	0
1	0	0	0	1	0
Σ	2	3	1	4	0

Table 5.4d

	A	B	D	C	E	Σ
	1	1	1	1	0	4
	1	1	1	0	0	3
	1	1	0	0	0	2
	1	0	0	0	0	1
Σ	4	3	2	1	0	10

Table 5.4e

	A	B	D	C	E	F_{II}
A	-	1	2	3	4	10
B		-	1	2	3	6
D			-	1	2	3
C				-	1	1
E					-	0

The
second
factor

Table 5.4f

Table 5.4. (continued)

	D	B	C	A	E
12	1	1	1	0	0
11	1	1	0	0	0
10	0	1	0	0	0
Σ	2	3	1	0	0

Table 5.4g

	B	D	C	E	A	Σ
	1	1	1	0	0	3
	1	1	0	0	0	2
	1	0	0	0	0	1
Σ	3	2	1	0	0	6

Table 5.4h

	B	D	C	E	A	F_{III}
B	-	1	2	3	3	9
D		-	1	2	2	5
C			-	1	1	2
E				-	0	0
A					-	0

The
third
factor

Table 5.4i

	D	B	C	A	E
13	1	0	1	0	0
Σ	1	0	1	0	0

Table 5.4j

	D	C	E	B	A	Σ
	1	1	0	0	0	2
Σ	1	1	0	0	0	2

Table 5.4k

	D	C	E	B	A	F_{IV}
D	-	0	1	1	1	3
C		-	1	1	1	3
E			-	0	0	0
B				-	0	0
A					-	0

The
fourth
factor

Table 5.4l

Table 5.5 Model C matrix of transformed scores from Table 5.1. Items and subject types were rearranged according to their respective marginals in descending order. Cell entries are ratios of Model A raw score matrix marginal products of the grand total of matrix elements.

		Items				
		D	B	C	A	F
	5	1.52	1.21	.91	.76	.60
	4	1.21	.97	.73	.60	.49
	6	1.21	.97	.73	.60	.49
	3	.91	.73	.55	.46	.36
	7	.91	.73	.55	.46	.36
	12	.91	.73	.55	.46	.36
Subject types	2	.60	.49	.36	.33	.24
	8	.60	.49	.36	.33	.24
	11	.60	.49	.36	.33	.24
	13	.60	.49	.36	.33	.24
	1	.33	.24	.18	.15	.12
	9	.33	.24	.18	.15	.12
	10	.33	.24	.18	.15	.12
	14	.00	.00	.00	.00	.00

Table 5.6 Model C dominance matrix for Factor I.

		Items				
		D	E	C	B	A
Subject types	5	1.67	1.33	1.00	.67	.33
	6	1.33	1.07	.80	.53	.21
	7	1.00	.80	.60	.40	.20
	8	.67	.53	.40	.21	.13
	9	.33	.21	.20	.13	.07
	14	.00	.00	.00	.00	.00

A counterparts (Tables 5.1 and 5.4a). Consider the geometric meaning of these matrix transformations. The data matrix is of order 14 x 5. The five elements of each row may be regarded as the Cartesian coordinates of a point in five dimensions. The rank of its monotonic transformation is four, with vanishing fifth order minors. Let us now take a close look at the transformed matrix for factor I. It can be seen that for any row i there exists a constant c_i such that $a_{ij} = d_i a_{1j}$. The same can be shown for columns. Each row (column) can be expressed linearly in terms of the other. This is not true for elements between particular factor matrices, which are linearly independent. According to Thurstone's theorem (1947, p. 282), the number of linearly independent factors represented by the intercorrelations of n tests is equal to the rank of their correlational matrix R, which is true for the data of Table 5.1, provided that a correlation matrix is reconceptualized as a dominance matrix.

In this sense the matrix in Table 5.7 is the matrix of order loadings, with column sums analogous to factor contributions and row marginals analogous to communalities. Note that the grand sum of 62 is equal to the number of one-zero changes in columns of our raw score matrix (i.e. in information theory terms, to the number of bits accounted for by items). Factor IV is due to the response pattern of one subject only

Table 5.7. Matrix of order loadings.

	F_I	F_{II}	F_{III}	F_{IV}	Σ
D	10	3	5	3	21
E	6	0	0	0	6
C	3	1	2	3	9
B	1	6	9	0	16
A	0	10	0	0	10
Σ	20	20	16	6	62

and can be considered residual. The latent structure of the data matrix, based on order loadings of Table 5.7 is presented in Figure 5.3.

We can now return to the original Conjunctive model of Figure 5.1. As compared with this model, the deterministic model of order analysis introduces several improvements. The major innovations are, (a) supplanting the estimate of lower bound dimensionality by determining the number of dimensions based on Thurstone's notion of the dimensionality indicator as matrix rank, and (b) presenting a computational algorithm which differs from the partially indeterminate procedure based on Bennett's theorem.

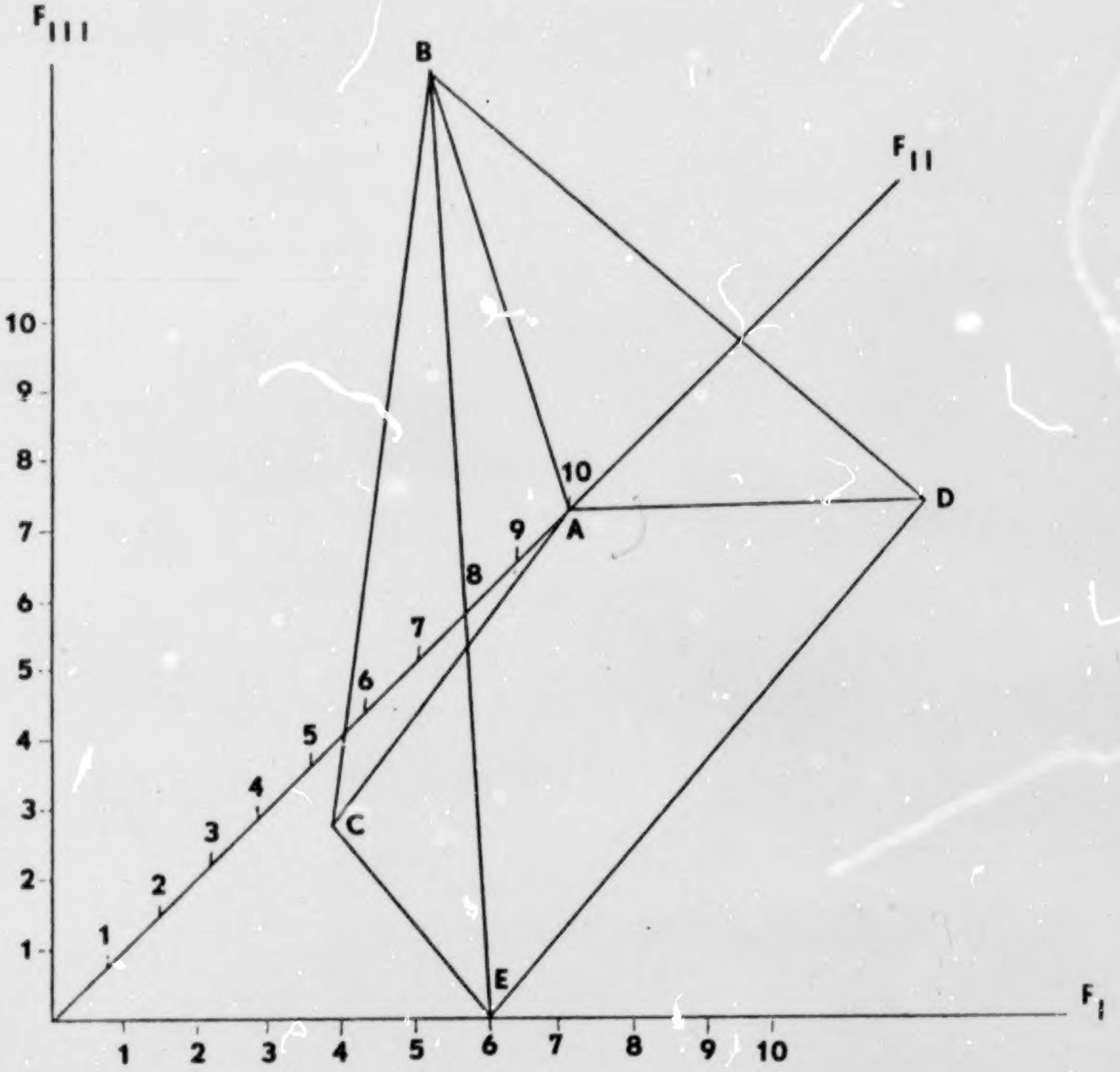

F_{III}
F_{II}
F_I
B
A
C
D
E
10
9
8
7
6
5
4
3
2
1
1
2
3
4
5
6
7
8
9
10
1
2
3
4
5
6
7
8
9
10

Figure 5.3. Latent structure of the data matrix D (Table 5.1) as isolated by the deterministic model of order analysis.

Chapter 6

THE PROBABILISTIC MODEL OF ORDER ANALYSIS

Probabilistic structure of Guttman-type data matrices.

Previously, we have obtained a matrix of magnitudes by consistently recording the frequencies of the negative implication function and its converse or by multiplying complementary parts of response vectors. If we let stand D_{rxq} for a binary data matrix of rxq order and $\bar{D}_{rxq}$ for its reflection (which can be conceptualized as the addition of the unit matrix U_{rxq} in the modulo two, i.e. $\bar{D}_{rxq} = D_{rxq} + U_{rxq}$ (mod 2)), then the matrix of magnitudes M_{qxq} can be written as

$$M_{qxq} = D'_{rxq} \cdot \bar{D}_{rxq} \qquad (6.1)$$

Suppose that an experiment will yield a data matrix as in Table 4.2, which is not a perfect Guttman-type data matrix, but reasonably close to it. This data matrix was obtained by administering a simple, twelve items scale (constructed to measure one attribute) to a group of fifteen subjects. To estimate its manifest structure, we have to obtain the matrix of magnitudes (Table 6.1) by premultiplying its reflection by its transpose, as by the use of equation 6.1. The entries m_{ij} of this matrix are frequencies of the negative implication and m_{ji} elements are frequencies of its converse; i.e. this

Table 6.1. The M_{qxq} matrix of magnitudes of changes for the data from Table 4.2. Elements of this matrix were obtained by premultiplication of the reflected data matrix by its transpose (Equation 6.1).

	A	B	C	D	E	F	G	H	I	J	K	L
A	0	1	4	1	5	7	0	4	1	0	2	0
B	7	0	6	2	7	13	2	7	4	2	4	0
C	5	1	0	2	4	8	0	3	3	1	2	0
D	5	0	5	0	5	11	2	6	3	1	3	0
E	5	1	3	1	0	7	2	4	3	1	2	0
F	0	0	0	0	0	0	0	0	0	0	0	0
G	6	2	5	4	8	13	0	8	5	1	5	0
H	3	0	1	1	3	6	1	0	3	1	3	0
I	3	1	4	1	5	9	1	6	0	0	1	0
J	6	2	6	3	6	13	1	9	4	0	4	0
K	4	1	3	1	4	9	1	6	1	0	0	0
L	8	2	7	4	8	15	2	9	6	2	6	0

matrix contains two opposite directions of order. If the data matrix contained random numbers, we would expect the frequencies of both negative implication and its converse to be approximately equal. Consider a data matrix consisting of two items A and B. Its corresponding matrix of magnitudes M_{qxq} would be in this case a fourfold table with frequencies m_{ij}, x_{jj}, m_{ji}, and x_{ii} starting in the first quadrant in clockwise direction. It is important to note that proportions p_a and p_b can differ only if m_{ij} differs from m_{ji}, since $p_a = (x_{ii} + m_{ij})$ and $p_b = (x_{ii} + m_{ji})$ have the common element x_{ii}. Frequency of the difference relation therefore equals $m_{ij} + m_{ji}$.

Assuming equiprobability of both directions we can test the significance of difference (for either direction) in two ways: (a) the binomial with mean equal to np and variance equal to npq; i.e. mean of $(m_{ij} + m_{ji})$ (.5) and variance of $(m_{ij} + m_{ji})$ (.5) (.5), or (b) with the normal curve approximation. The latter is accomplished by expressing the difference relation as a deviation from the mean and divided by its standard deviation, which gives the critical ratio:

$$\frac{m_{ij} - (m_{ij} + m_{ji})\,(.5)}{[(m_{ij} + m_{ji})\,(.5)\,(.5)]^{\frac{1}{2}}} = \frac{m_{ij} - .5m_{ij} - .5m_{ji}}{.5(m_{ij} + m_{ji})^{\frac{1}{2}}} = \qquad (6.2)$$

$$\frac{m_{ij} - m_{ji}}{(m_{ij} + m_{ji})^{\frac{1}{2}}}$$

Equation 6.3 is similar to the McNemar equation (5.3) for a critical ratio of nonindependent proportions (McNemar, 1969, pp. 54-58).

The matrix of magnitudes of changes can be partitioned symmetrically along the diagonal into the upper triangular matrix M^{10}_{qxq} of negative implication frequencies and lower triangular matrix M^{01}_{qxq} of frequencies of its converse:

$$M_{qxq} = [M^{10}_{qxq} \mid M^{01}_{qxq}] \qquad (6.3)$$

The matrix of z ratios then can be written as:

$$Z_{qxq} = \frac{M^{10}_{qxq} - M^{01'}_{qxq}}{(M^{10}_{qxq} + M^{01}_{qxq})^{\frac{1}{2}}} + \frac{M^{01}_{qxq} - M^{10'}_{qxq}}{(M^{01}_{qxq} + M^{01}_{qxq})^{\frac{1}{2}}} \qquad (6.4)$$

This matrix is listed in Table 6.2 for our example.

A series of the order pattern matrices ${}^{p}S$ (dependent on the lower bound (z_p, $0 \leq 1 \leq +\infty$) of the corresponding confidence interval) can be constructed for every z_p significance level. The entries of order pattern matrices s_{ij} are determined by $s_{ij} = m_{ij}$ and $s_{ji} = m_{ji}$ if the z_{ij} value is inside its confidence interval. The relationship $s_{ij} = s_{ji} = 0$ is true if the z_{ij} value is outside its confidence interval. The lower bound of the confidence interval is specified by the researcher. Successive scanning of the whole positive interval of a_{ij} values

Table 6.2. The Z_{qxq} matrix of z ratios for the dominance matrix M_{qxq} (Table 6.1).

	A	B	C	D	E	F	G	H	I	J	K	L
A	0.00	-2.12	-0.33	-1.64	0.00	2.65	-2.42	0.38	-1.00	-2.45	-0.82	-2.83
B	2.12	0.00	1.89	1.42	2.12	3.60	0.00	2.65	1.34	0.00	1.34	1.42
C	0.33	-1.89	0.00	-1.13	0.38	2.83	-2.23	1.00	-0.38	-1.89	-0.45	-2.65
D	1.64	-1.42	1.13	0.00	1.64	3.32	-0.82	1.89	1.00	-1.00	1.00	-2.00
E	0.00	-2.12	-0.38	-1.64	0.00	2.65	-1.90	0.38	-0.17	-1.89	-0.82	-2.83
F	-2.65	-3.60	-2.83	-3.32	-2.65	0.00	-3.60	-2.45	-3.00	-3.60	-3.00	-3.89
G	2.45	0.00	2.23	0.82	1.90	3.60	0.00	2.34	1.64	0.00	1.64	-1.42
H	-0.38	-2.65	-1.00	-1.89	-0.38	2.45	-2.34	0.00	-1.00	-2.53	-1.00	-3.00
I	1.00	-1.34	0.38	-1.00	0.71	3.00	-1.64	1.00	0.00	0.00	0.00	-2.45
J	2.45	0.00	1.89	1.00	1.89	3.60	0.00	2.53	2.00	0.00	2.00	-1.42
K	0.82	-1.34	0.42	-1.00	0.82	3.00	-1.64	1.00	0.00	0.00	0.00	-2.45
L	2.83	1.42	2.65	2.00	2.82	3.89	1.42	3.00	2.45	1.42	2.45	0.00

is also possible. Thus,

$$^{P}S_{q\times q} = M_{q\times q}; \text{ where } s_{ij} \begin{array}{l} = m_{ij} \text{ if } z_{ij} \geq z_p \\ = 0 \text{ if } z_{ij} < z_p \end{array} \qquad (6.5)$$

By diagonal symmetric partitioning of the order pattern matrix ^{P}S,

$$^{P}S_{q\times q} = [{}^{P}S^{10}_{q\times q} \mid {}^{P}S^{01}_{q\times q}] \qquad (6.6)$$

and by reduction to the unit space, our relational space can be reduced into a unit space as

$$^{P}P_{q\times q} = \frac{{}^{P}S^{10}_{q\times q} - S^{01'}_{q\times q}}{N} + \frac{{}^{P}S^{01}_{q\times q} - S^{10'}_{q\times q}}{N} \qquad (6.7)$$

To facilitate understanding of the described matrix operations, the matrices of distances are recorded in Tables 6.3, 6.4, 6.5 for significance levels $z_1 \geq 2.58$, $z_2 \geq 1.00$, and $z_3 > 0.00$.

Unidimensional components of the connected graphs and their extraction.

The relevance of the series of these metric relational spaces for Guttman scaling can be best depicted by the means of tree diagrams. A tree is a connected graph which can be used to plot the logical possibilities of a sequence of events where each event can occur in a finite number of ways. There are

Table 6.3. The matrix $p_{qxq}^{z \geq 2.58}$ of the structure of the test space. Elements of this matrix were obtained from the dominance matrix M_{qxq} by matrix transformations described by equations 6.3 through 6.7. They are spatial generalizations of the sample space of the test into the unit population space.

	A	B	C	D	E	F	G	H	I	J	K	L
A	-	-	-	-	-	.467	-	-	-	-	-	-
B	-	-	-	-	-	.867	-	.476	-	-	-	-
C	-	-	-	-	-	.533	-	-	-	-	-	-
D	-	-	-	-	-	.733	-	-	-	-	-	-
E	-	-	-	-	-	.467	-	-	-	-	-	-
F	-	-	-	-	-	-	-	-	-	-	-	-
G	-	-	-	-	-	.867	-	-	-	-	-	-
H	-	-	-	-	-	-	-	-	-	-	-	-
I	-	-	-	-	-	.600	-	-	-	-	-	-
J	-	-	-	-	-	.867	-	-	-	-	-	-
K	-	-	-	-	-	.600	-	-	-	-	-	-
L	.533	-	4.67	-	.533	1.000	-	.600	-	-	-	-

Table 6.4. The matrix $p_{qxq}^{z \geq 1.00}$ of the structure of the test space. Elements of this matrix were obtained from the dominance matrix M_{qxq} by matrix transformations described by equations 6.3 through 6.7. These transformations allow for the spatial generalization of the sample test space into the unit population space.

	A	B	C	D	E	F	G	H	I	J	K	L
A	-	-	-	-	-	.467	-	-	-	-	-	-
B	.400	-	.333	.133	.400	.867	-	.476	.200	-	.200	-
C	-	-	-	-	-	.533	-	.133	-	-	-	-
D	.267	-	.200	-	.267	.733	-	.333	.133	-	.133	-
E	-	-	-	-	-	.467	-	-	-	-	-	-
F	-	-	-	-	-	-	-	-	-	-	-	-
G	.400	-	.333	-	-	-	-	-	-	-	-	-
H	-	-	-	-	-	.400	-	-	-	-	-	-
I	.133	-	-	-	-	.600	-	.200	-	-	-	-
J	.400	-	.333	.133	.333	.867	-	.533	.267	-	.267	-
K	-	-	-	-	-	.600	-	.200	-	-	-	-
L	.533	.133	.467	.267	.533	1.000	.133	.600	.400	.133	.400	-

Table 6.5. The matrix $P^{z > 0.00}_{qxq}$ of the structure of the test space. Elements of this matrix are obtained from the dominance matrix M_{qxq} by matrix transformations described by equations 6.3 through 6.7.

	A	B	C	D	E	F	G	H	I	J	K	L
A	-	-	-	-	-	.467	-	.067	-	-	-	-
B	.400	-	.333	.133	.400	.867	-	.467	.200	-	.200	-
C	.067	-	-	-	.067	.533	-	.133	-	-	-	-
D	.267	-	.200	-	.267	.733	-	.333	.133	-	.133	-
E	.000	-	-	-	-	.467	-	.067	-	-	-	-
F	-	-	-	-	-	-	-	-	-	-	-	-
G	.400	.000	.333	.133	.400	.867	-	.476	.267	-	.267	-
H	-	-	-	-	-	.400	-	-	-	-	-	-
I	.133	-	.067	-	.133	.600	-	.200	-	-	-	-
J	.400	.000	.333	.133	.333	.867	.000	.533	.267	-	.267	-
K	.133	-	.067	-	.133	.600	-	.200	.000	-	-	-
L	.533	.133	.467	.267	.533	1.000	.133	.600	.400	.133	.400	-

various types of trees. Relational data from our matrix of distances ${}^{P}P_{qxq}$ can be imagined in the general case as a plane rooted tree whose branches are the best approximation of the Guttman-type data matrices. The degree of approximation is a direct function of the z_p parameter of our model.

There are several ways to extract the branches of connected, asymmetrical, and transitive relations from the ${}^{P}P_{rxr}$ matrix. One way is to consider transitional probabilities at every node of the tree, which was the approach we adopted in the computerized version of this model (Krus, 1973). Recently we have adapted a different approach, following the coding algorithms for unlabeled trees. This algorithm was developed in graph theory by DeBruijn and Morselt (1967), Thalwitzer (1968) and Read (1972), as a consequence of Cayley's (1875) ideas. The coding procedure is to map a graph into a numerical code so that decoding will produce a tree which is isomorphic with the coded tree.

Read's (1972) coding procedure is quite simple and can be best understood by coding the actual tree diagram (Figure 6.1). The coding is by binary integers C_{gh}, where g is the number of tree nodes and h the number of tree roots. The tree in Figure 6.1 has two roots at nodes H and F. The strings of binary codes can be written separately for every root as pyramidal catenations of modulo two integers. Thus, for the root H,

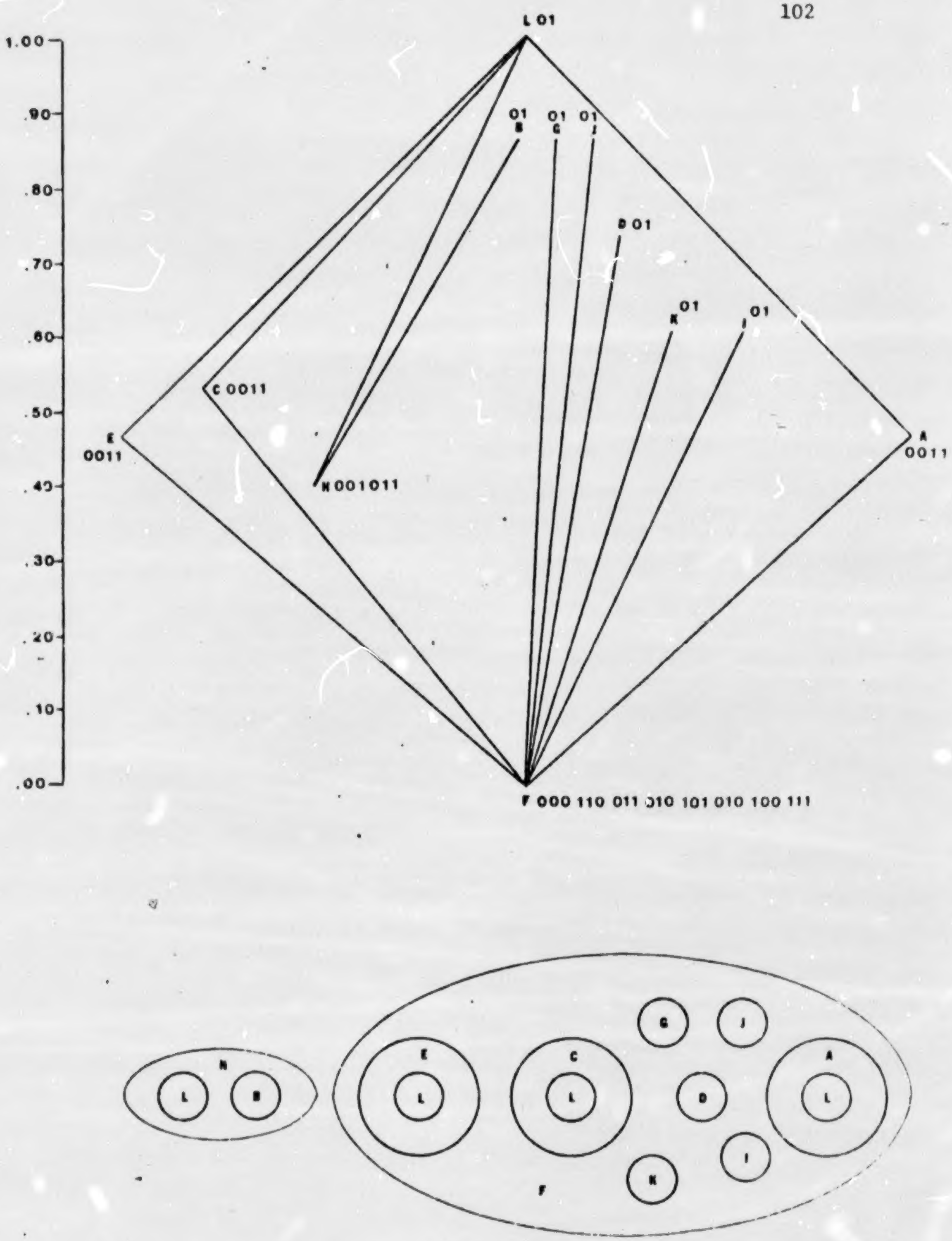

1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
L 01
01 B
01 G
01 J
D 01
K 01
I 01
C 0011
E 0011
A 0011
H 001011
F 000 110 011 010 101 010 100 111
H
L
B
E
L
C
L
G
J
D
A
L
I
K
F

Figure 6.1. Sample manifest structure of the generalized ($z_p \geq 2.58$) dominance matrix in a unit space.

the code is 001011 or 13 octal and for root F, the binary code is 000 110 011 010 101 010 100 111 or 6, 325, 247 octal. The decoding procedure is done by means of closed curves. It is possible to imagine the code for every root as a closed system of left (0) and right (1) parentheses (Read, 1972, p. 161); i.e. for the H root:

(()())

and for the F root:

((())(())()()()()()(()))

The bottom part of Figure 6.1 was constructed by enclosing a node into a circle or elipse according to its position in the system of parentheses. Closed curves represent tree nodes. Adjacent edges are defined by the immediate inclusion of one curve inside another. It is possible to reconstruct the original tree from this type of code either by this graphic procedure or by the matrix algorithm. To accomplish the latter, a matrix is produced with elements symbolizing those pairs of nodes that make up the edges of the tree.

Inspection of the bottom part of Figure 6.1 reveals that node L is present in more than one closed curve configuration and is, in this sense, redundant. To remove this redundancy, the following procedure was developed.

Consider the graph in Figure 6.2, constructed for the same data by arbitrarily setting the value of $z \leq 1.0$.

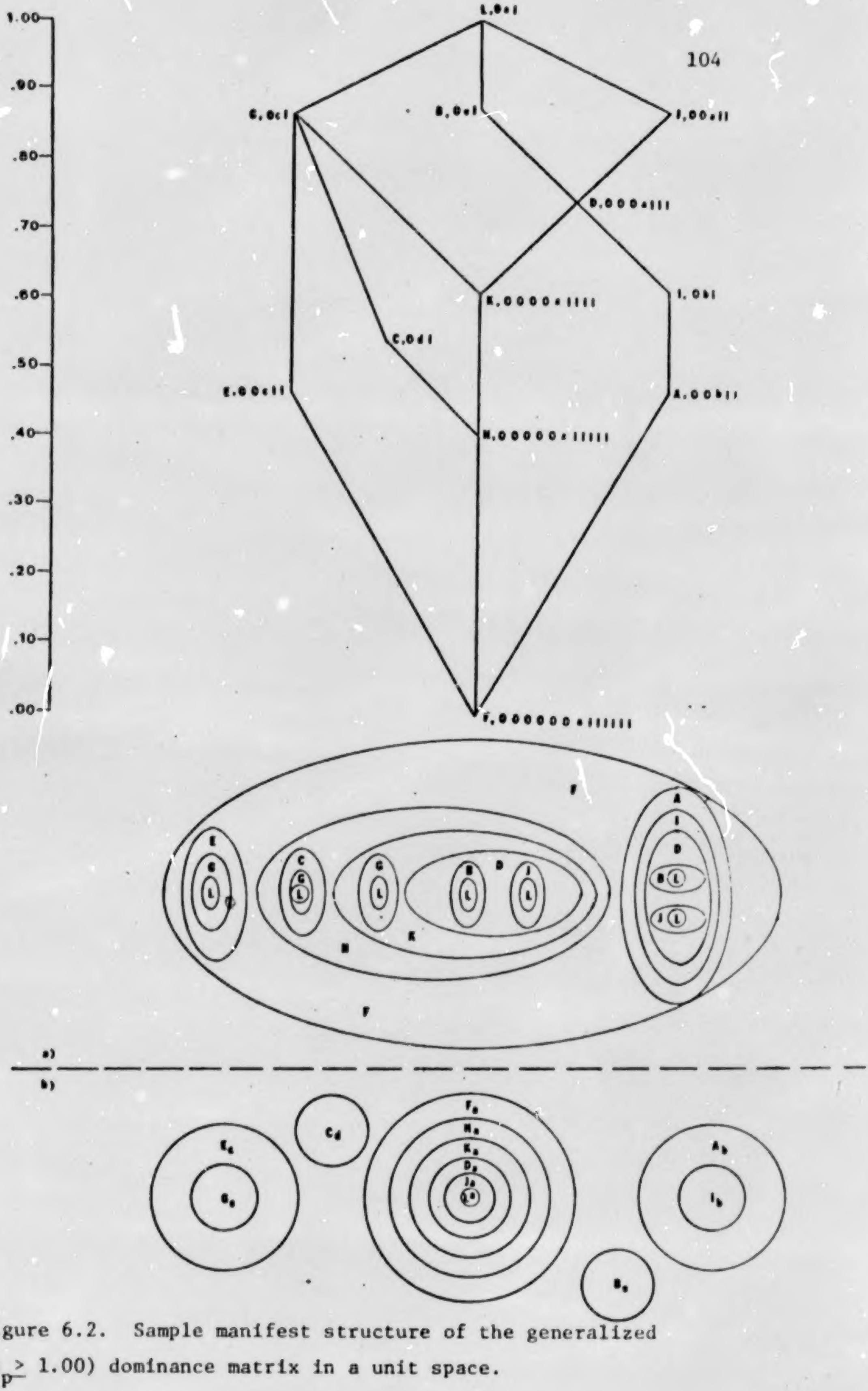

1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
L,0a1
G,0c1
B,0e1
J,00a11
D,000a111
K,0000a1111
I,0b1
C,0d1
E,00c11
A,00b11
H,00000a11111
F,000000a111111
F
E
G
L
C
G
L
G
L
B
L
D
J
L
K
H
A
I
D
B
L
J
L
F
a)
b)
F_a
H_a
K_a
D_a
J_a
L_a
C_d
E_c
G_c
A_b
I_b
B_e

Figure 6.2. Sample manifest structure of the generalized ($z_p \geq 1.00$) dominance matrix in a unit space.

Read's code can be computed for every node as L(01), G(0011), B(0011), ... , F(000011100000111000110001100111110000011001111111). The octal code for the node F would be $10^{12} \cdot 7_8$. For our purposes a simpler coding can be adopted by defining an index K_{gh} as a separate sum of binary integer codes C_{gh} for every tree node. Values of the K_{gh} index are reported for the illustrative example in Table 6.6. In step 1, the sums of Read's C_{gh} codes were obtained separately for every node g and arranged in descending order in step 2. Concatenation of tree nodes in this order (and according to their adjacent edges) will yield their roots h (in our case a,b,c,d, and e), as recorded in step 3 of Table 6.6.

The redundant projection of the logic diagram, coded by the C_{gh} index is depicted in Figure 6.2a (the bottom part). Projection of nonredundant configurations, as coded by the K_{gh} index, are recorded in Figure 6.2b.

As mentioned earlier, every magnitude in the matrix M_{rxr} was obtained by recording a value 1 for every ("10") and("01") relation between the data matrix elements. This dominance-type matrix preserves orders. This order preserving property was maintained through subsequent transformations and formed a basis for a particular tree construction. The resulting tree code also preserves this property. The breaks between the substrings of zeros and ones represent the informational

Table 6.6. Values of the K_{gh} index and roots of the logic diagram drawn in Figure 6.2.

K_{gh}	K_{gh}	TREE ROOTS
A = 7	F = 23	a
B = 2	H = 12	a
C = 3	K = 8	a
D = 5	A = 7	b
E = 3	I = 6	b
F = 23	D = 5	a
G = 2	C = 3	c
H = 12	E = 3	d
I = 6	G = 2	d
J = 2	B = 2	e
K = 8	J = 2	a
L = 1	L = 1	a
(Step 1)	(Step 2)	(Step 3)

one-zero breaks between the elements of the data matrix ordered in a certain direction and dimension. By considering the tree nodes for every separate area and by writing the reordered n-tuples of the original data matrix in a sequence of the tree codes, (as e.g. for the tree branch *a* from Figure 6.2):

L	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
J	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1
D	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0
I	1	1	1	1	1	1	1	1	1	0	0	0	0	1	0
A	1	1	1	1	0	0	1	0	0	1	0	0	0	1	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

an approximation to the Guttman-type data matrix is obtained. The degree of approximation is a function of z_p.

If we continue to lower the z_p values, the "width" of the tree diagrams (which is itself an artifact of planary projection of the tree branches from the multi-dimensional space) will diminish. The multifactorial composition of the test space will be forced by our probabilistic model into a single dimension, i.e. at a certain value of the z_p parameter, the test space will collapse into one dimension. In our example, during the scanning of the z interval, it happened when the lower bound of the z_p interval approached zero (Figure 6.3). At this point, the tree diagram changed into a linear, unidimensional scale identical with the Guttman scale (as constructed in Table 6.7) and the graphic representation of both its codes became concentric (bottom part of the Figure 6.3). This "point of unidimensionality" is the function of the dimensionality of the data matrix. If ordering in one dimension is possible, it will be located in the positive interval of the z scale.

If the data matrix contains more than one distinct order, any attempt to construct a unidimensional scale would mean that our probability of being right is less than .50; i.e., we would be constructing a model which would be more probably incorrect than correct. In this limiting case, the use of the multidimensional model becomes mandatory. The reasonable requirement is to stay within the limits of the classic one or five per cent significance levels, which is the practice meticulously observed in significance testing. The majority of multivariate analysis models do not permit this direct observation of the confidence intervals for their dimensionality estimates. According to our preliminary comparisons of our model with the factor analytic methods, the factorial and order analytic structures frequently converge in the $0 < z_p < 1.00$ intervals, which can account for their instability.

When the conversion of the data matrix into one dimension would be highly unstable, it is possible to partition it into

Table 6.7. The scalogram analysis of the data matrix from Table 4.2. The Guttman scale was constructed by setting cutting points in such a way, that it would be necessary to reflect the minimal number of raw data matrix elements (from one to zero to the left of the cutting points and from zero to one on their right sides). The unit scale was obtained by dividing each step of Guttman's scale by the number of subjects.

		Subject types															G-scale	
		4	10	9	8	5	11	12	15	14	13	7	2	3	1	6	Raw	Unit
Items	L	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	1.000
	J	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	13	.867
	G	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	13	.867
	B	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	13	.867
	D	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	11	.733
	K	0	0	0	0	1	0	1	1	1	0	1	1	1	1	1	9	.600
	I	0	1	0	0	0	0	1	1	1	0	1	1	1	1	1	9	.600
	C	0	0	1	0	1	0	0	1	0	1	1	0	1	1	1	8	.533
	E	0	0	0	1	1	0	1	0	0	1	1	1	0	0	1	7	.467
	A	0	1	0	0	0	1	0	0	1	0	0	1	1	1	1	7	.467
	H	0	0	1	1	0	0	0	0	0	1	0	0	1	1	1	6	.400
	F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	.000

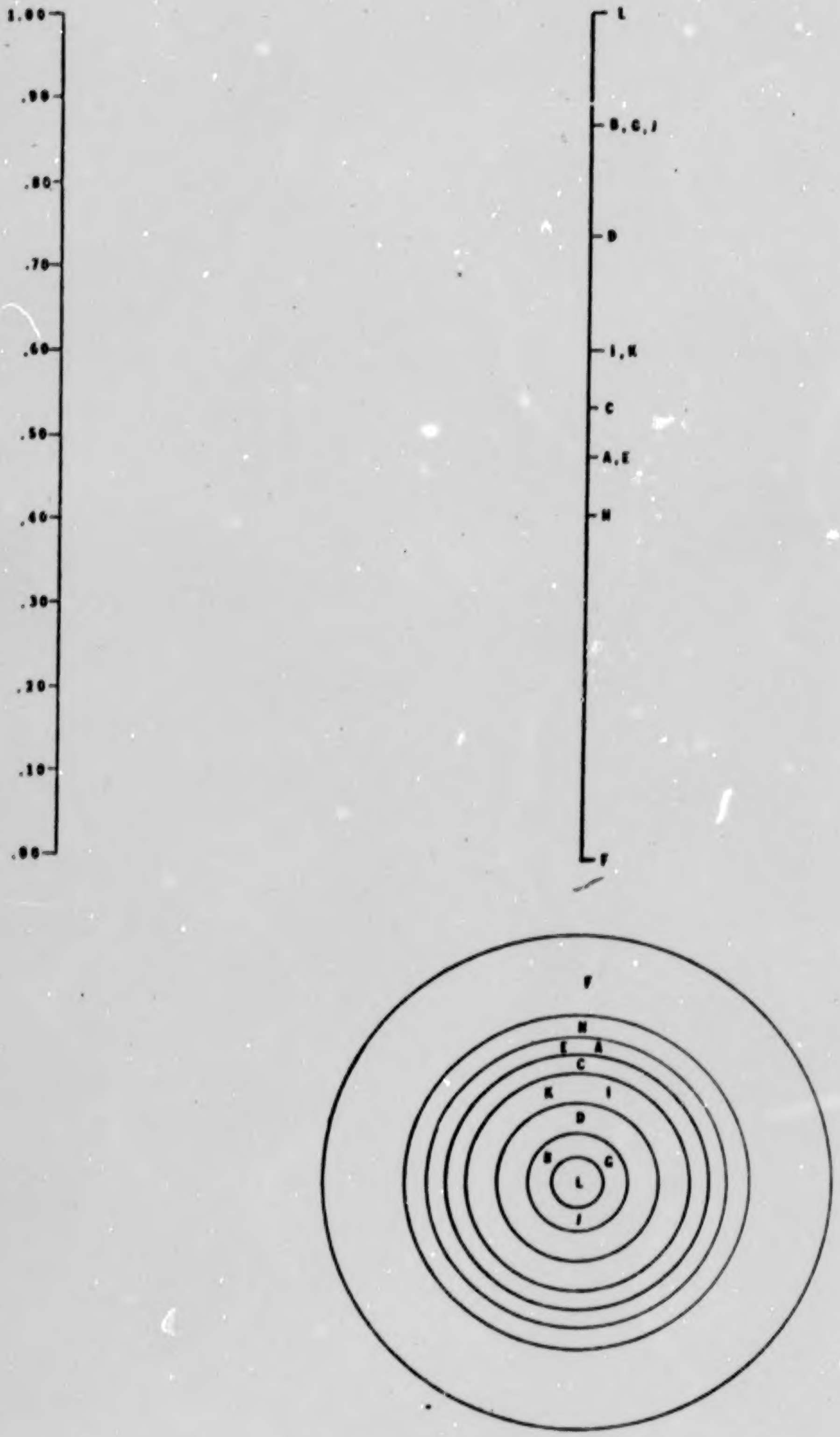

1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
L
B,G,J
D
I,K
C
A,E
H
F
F
H
E
A
C
K
I
D
B
G
L
J

Figure 6.3. Sample manifest structure of the dominance matrix in a unit space ($z_p > 0.00$).

a set of mutually orthogonal submatrices with a potential for the unidimensional direction of order. We will describe this transition of the data matrix into the structured supermatrix in the following section.

Latent dimensions of the data matrix.

The previously stated goal of order analysis is the isolation of a data matrix's manifest and latent attributes and entities. Manifest attributes or entities were conceptualized in terms of subsets of ordered n-tuples and described in terms of asymmetrical, transitive, and connected relations between the elements of each subset. The data matrix (if we limit our discussion to the test-persons plane of Cattell's covariation chart (Cattell, 1946, pp. 96-101)) contains information about the latent classes of subjects (entities) and items (attributes). Each latent class of subjects can be isolated if the relations between its members comply with the criteria set up as the necessary and sufficient conditions for potential unidimensional order. Partitioning the data matrix for its entities and attributes will result in submatrices with potential for ordering the opposite grouping type.

Thus the integrated analysis of the information structure of the data matrix in the test-persons plane would necessitate the extraction from the data matrix D_{rxq} the matrix of magnitudes of hypothetical observed changes M_{rxr} in the subject's direction of order, i.e.

$$M_{rxr} = D_{rxq} \, . \, \overline{D}'_{rxq} \tag{6.8}$$

with the corresponding diagonal partitioning

$$M_{rxr} = [M^{10}_{rxr} \mid M^{01}_{rxr}] \tag{6.9}$$

The formation of subject's typal groups at a specific z_p level is done by operations analogous to the separation of the tree branches of the item space; the relations analyzed are those between subjects. Hence,

$$Z_{rxr} = \frac{M^{10}_{rxr} - M^{01'}_{rxr}}{(M^{10}_{rxr} + M^{01}_{rxr})^{\frac{1}{2}}} + \frac{M^{01}_{rxr} - M^{10'}_{rxr}}{(M^{10}_{rxr} + M^{01}_{rxr})^{\frac{1}{2}}} \tag{6.10}$$

and

$$^{p}S_{rxr} = M_{rxr}; \text{ where } s_{ij} \begin{cases} = m_{ij} \text{ if } z_{ij} \geq z_p \\ = 0 \text{ if } z_{ij} < z_p \end{cases} \tag{6.11}$$

Triangulation of the $^{p}S_{rxr}$ matrix and subsequent reduction of the reordered M_{rxr} matrix into the unit space is also analogous to our previously described algorithm, but in this case matrix operations are done on the transposed matrices:

$$^{P}S_{rxr} = [S^{10}_{rxr} \mid S^{01}_{rxr}] \qquad (6.12)$$

and

$$^{P}P_{rxr} = [\tfrac{1}{N}\,(S^{10}_{rxr} - S^{01'}_{rxr} + S^{01}_{rxr} - S^{10'}_{rxr})] \quad . \qquad (6.13)$$

By nonredundant coding of the resulting trees, the root codes are obtained and D_{rxq} matrix is regrouped and partitioned according to their descending order,

$$D_{rxq} = \left[\begin{array}{c} D^{d1}_{n1xq} \\ \hline D^{d2}_{n2xq} \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline D^{dk}_{ngxq} \end{array}\right] \qquad (6.14)$$

where n_1, n_2, ... n_g designate the number of nodes in the separated and replanted branches of the original tree with the number of new roots d_1, d_2, ... d_k dependent on the z_p level. These operations on the intersubject relations of the data matrix aim at the isolation of item's latent attributes. Attributes underlying the observed variation between test items are thought of as ordered variables. Despite the fact that this order is inferred by the formal matrix operations, it was imposed on the data matrix by subjects, whose judging operations the data matrix reflects. The variation between subjects corresponds to the interindividual differences with respect to an attribute if the potential for unidimensional order is present in a particular subgroup's responses. The potential unidimensional order of a different subgroup reflects a different attribute. The partitioned data supermatrix can be further analyzed by observing the magnitudes of interitem relations; as

$$M^{d}_{qxq} = \left[\begin{array}{c} D^{d1'}_{n1xq} \cdot \bar{D}^{d1}_{n1xq} \\ \hline D^{d2'}_{n2xq} \cdot \bar{D}^{d2}_{n2xq} \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline D^{dk'}_{ngxq} \cdot \bar{D}^{dk}_{ngxq} \end{array}\right] \qquad (6.15)$$

The resulting system of M^{d}_{qxq} matrices is redundant in the sense of possible repetitions of the hypothetical simultaneous observation of a difference with respect to a property measured. The isolation of this redundancy is desirable and theoretically could be accomplished by the separate computation of the major product of response vector with its reflection for every subject. A dominance matrix could be constructed by recording one bit of information for every observation, irrespective of the number of subjects by which it was observed. This type of redundancy removal would be appropriate for small data matrices in the deterministic model.

In the probabilistic formulation, the report of a single individual is not liable to verification in a formal sense. The inclusion of each particular bit of information into the final scales should, therefore, be conditionally dependent on the degree of certainty that the reported change is real. This is analogous to the distinction between the common and specific variance of the factor analytic model. When the observed and reported information should be called common is, of course, the question. Are changes common if reported by two subjects, by all subjects, or by all but one?

The probabilistic solution could be to define the isolated magnitudes of changes as common if shared by more than 95% of the total sample with the remaining percentage of magnitudes defined as specific. In matrix notation, for every isolated dimension d_k; where k=1, 2, ... k, ..., m, we can compute from dimension-specific dominance matrices M^{d1}_{qxq}, M^{d2}_{qxq}, ..., M^{dk}_{qxq}, ... M^{dm}_{qxq}, their counterparts, as

$$Z^{d1}_{qxq},\ Z^{d2}_{zqz},\ \ldots,\ Z^{dk}_{qxq},\ \ldots,\ Z^{dm}_{qxq} \qquad (6.16)$$

where Z_{qxq} is defined as in equation 6.5.

The matrix F_{qxm}, describing the latent structure of attributes, can be constructed from the row marginal sums of reduced matrices of magnitudes in every dimension. The reduced matrices of magnitudes can be written as

$$_{d_k}G^p_{q\times q} = {_{d_1}G^p_{q\times q}}, {_{d_2}G^p_{q\times q}}, \ldots, {_{d_k}G^p_{q\times q}}, \ldots, {_{d_m}G^p_{q\times q}};$$

$$\text{where } g_{ij} \begin{cases} = m_{ij} \text{ if } z_{ij} \geq z_p \\ = 0 \text{ if } a_{ij} < z_p \end{cases} \qquad (6.17)$$

Other transformations are possible, as e.g.

$$g_{ij} \begin{cases} = 1 \text{ if } z_{ij} \geq z_p \\ = 0 \text{ if } a_{ij} < z_p \end{cases} \qquad (6.18)$$

or

$$g_{ij} \begin{cases} = a_{ij} \text{ if } z_{ij} \geq z_p \\ = 0 \text{ if } a_{ij} < z_p \end{cases}$$

which have interesting scale properties not to be discussed here.

Chapter 7

VALIDITY OF ORDER ANALYSIS

Prestructured data (monotonic transformations).

Methods of multidimensional analysis should reflect the geometry of real objects in our terrestrial space. The forms provided by the earth are in a sense limited and simpler than ideational or other types of abstract spaces. Correct reflection of forms suggested by ordinary space is a fundamental validity requirement for any method of dimensional analysis. This type of validation is possible, if a set of simple physical objects with well-known relations between its dimensions, is analyzed.

A collection of eight boxes was assembled (Table 7.1) and their respective measures scrambled by twelve monotonic transformations (nine of them borrowed from Thurstone, 1947, p. 142). The binary data matrix created by these operations was dichotomized at the median, and order analyzed by the deterministic model. The hypothesis of this validation experiment was that order analysis will recover the original data matrix dimensions of length (x), width (y), and height (z). This hypothesis was verified as reported in Table 7.2. Order loadings are reported in bits (as rotated by Varimax). The highest order loadings for every variable delimited the first dimension as a length

Table 7.1. Measurements of length (x), width (y) and height (z) of a collection of eight boxes.

	x	y	z
Box 1	5	8	6
Box 2	7	2	9
Box 3	2	5	2
Box 4	6	3	4
Box 5	9	6	7
Box 7	3	4	8
Box 8	8	7	3

Table 7.2. Order analysis (deterministic model) of monotonically transformed and dichotomized data from Table 4.1.

Variables	X Dimension	Y Dimension	Z Dimension
X^2	12.40	1.80	1.80
Y^2	1.84	12.38	1.82
Z^2	1.83	1.78	12.40
log X	12.40	1.77	1.80
log Y	1.84	12.38	1.82
log Z	1.83	1.78	12.39
e^x	12.40	1.77	1.79
e^y	1.84	12.38	1.82
e^z	1.83	1.78	12.39
$5x + 3$	12.40	1.77	1.79
$7y + 8$	1.84	12.38	1.82
$2z + 5$	1.83	1.78	12.39

factor (x), the second dimension was determined by variables derived from height measures (z) and the third dimension by measures pertaining to width (y).

Distorted data (nonmonotonic transformations).

Thurstone's experiment with nonmonotonically distorted measures of real cubic objects known as "the box problem" (Thurstone, 1947, pp. 140-143) recently attracted attention, after it was realized that "all the common analytic and numerical methods fail to locate the clear simple structure when applied to it" (Cureton and Mulaik, 1971, p. 375). This particular experiment is also a validity-type test. Moreover, its stringent requirements can facilitate additional insights into properties of validated techniques.

Thurstone's original data was reconstructed from Tables 4 and 5 (Thurstone, 1947, pp. 141-142). The use of logical constants by order analysis necessitated conversion of Thurstone's data matrix into a binary matrix by dichotomization at the mean. Comparisons of rotated (Varimax) factor and order analytic solutions can be made from Tables 7.3 and 7.4. Principal component factor analysis was done on a matrix of PHI correlations and order analysis of the dominance matrix was at 1.64 z level. For factor analysis, the retained eigenvalues were 9.81, 3.58, 2.76 and 1.17. Order analysis retained four factors, and rejected two factors with extraction indexes

Table 7.3. Rotated matrix of factor loadings for Thurstone's "box problem". Rotation was by Varimax with Kaiser's normalization.

Variables	(X) Factor 1	(Y) Factor 2	(Z) Factor 3	Factor 4
X^2	.98	.10	.04	-.02
Y^2	.13	.92	-.01	.03
Z^2	.05	.10	.35	.92
XY	.39	.86	.19	.05
XZ	.26	.08	.82	.23
YZ	.00	.27	.85	.27
$\sqrt{X^2 + Y^2}$	.70	.60	.06	-.01
$\sqrt{X^2 + Z^2}$	.82	.18	.17	.31
$\sqrt{Y^2 + Z^2}$	.10	.73	.24	.51
2X + 2Y	.39	.86	.17	.05
2X + 2z	.52	.14	.57	.21
2Y + 2Z	.04	.63	.58	.22
log X	.98	.10	.04	-.02
log Y	-.09	.55	.38	.21
log Z	.10	-.03	.84	.12
XYZ	.21	.38	.65	.46
$\sqrt{X^2 + Y^2 + Z^2}$	.52	.65	.27	.18
e^x	.83	.23	.19	-.02
e^y	.13	.92	-.01	.03
e^z	.05	.11	.35	.92

Table 7.4. Matrix of order loadings for Thurstone's "box problem." This solution was rotated by Varimax with Kaiser's normalization. Order loadings are reported in information bits.

Variables	(x) Dimension 1	(y) Dimension 2	(z) Dimension 3	Dimension 4
x^2	27.00	3.06	5.54	-.27
y^2	2.87	20.13	4.28	2.86
z^2	1.04	-.18	9.59	9.12
xy	5.51	19.52	8.47	1.72
xz	16.53	1.20	40.42	1.75
yz	-1.31	17.94	29.08	11.42
$x^2 + y^2$	23.84	22.05	10.36	1.86
$x^2 + z^2$	26.16	4.56	12.20	1.16
$y^2 + z^2$	2.30	17.10	13.14	11.67
2x + 2y	5.51	19.52	8.47	1.72
2x + 2z	33.76	5.60	34.71	6.32
2y + 2z	1.88	29.64	30.10	13.90
log x	26.99	3.06	5.54	-.27
log y	1.97	25.70	32.06	38.24
log z	13.16	13.17	58.38	6.10
xyz	1.84	5.12	14.83	7.77
$x^2+y^2+z^2$	11.17	17.55	20.38	-.30
e^x	14.68	-.08	9.62	-3.54
e^y	2.87	20.13	4.28	2.86
e^z	-1.04	-.18	9.59	9.12

lower than .05. (The meaning of the extraction index will be discussed later.) As compared with Thurstone's (1947, p. 371, Table 2) original solution of the box problem (centroid), the principal components solution (Table 7.3) and order analysis (Table 7.4) isolated an additional dimension. In the case of factor analysis this additional dimension was probably generated by the point-biserial correlation coefficients. (This guess could be substantiated if a matrix of tetrachorics were analyzed and both solutions compared.) In the case of order analysis, The solution was probably influenced by a nonstandardized dominance matrix. Again, this could be tested if double standardization routines were available for the order analysis program.

Inspection of both factor and order loadings shows that both methods returned correct structures. Exceptions are loadings for variables z^2 and e^z in the case of factor analysis and variable log y loading high on the fourth dimension in the structure returned by order analysis. When all six dimensions isolated by order analysis were rotated and highest variable loadings separately highlighted for the first three and second three dimensions, a structure $x_1, y_1, z_1 \mid x_2, y_2, z_2$ emerged.

Random data.

The tendency of factor analytic models to analyze patterns of random variation is well known. Thus Guilford's (1967) classic article "When Not to Factor Analyze" cautions that "statistical operations do not compensate for carelessness in making observations" (Guilford, 1967, p. 310). This caveat is based on repeated observations of this fallacy; were factor analysis to reflect random structures as such, this type of warning would be unnecessary. It is unfortunately true that factor analytic methods do analyze random variation present in the data matrix and that, from inspection of factorially derived structures alone, it is impossible to ascertain the degree of randomness of the data.

The z_p parameter order analysis provides the technical means for alteration of sensitivity to random variation. As may be recalled, order analysis at the specific z_p level selects from the data matrix elements only those orders with probability of particular direction equal to or greater than the preset probability z_p. Thus order analysis at the 1.64 z level analyzes only orders in which the probability of a particular direction p is equal to or greater than .95. The decision to accept vectors of less certain determination is simulated by the lowering of the z_p level. Thus order analysis at the .01 z level (equivalent to .50 p level), accepts as real those orders with probability p greater than .50, i.e. all potential orders within the .01 to + ∞ z-interval which is equivalent to the .50 to + 1.00 p interval.

If a data structure is highly organized (i.e., nonrandom), the majority of ordered vectors would be expected to lie in the upper part of the z interval and the resultant structure would be a correct reflection of the real structure of the data matrix. If the data matrix contains substantial amounts of random variation, the majority of potential orders will be in the lower part of the z interval and in all likelihood will be unstable.

An indirect test of this hypothesis would be comparison of structures isolated by different models from random data matrices with simultaneous variation of the order analysis z parameter. Using random data matrices, it was expected that the factor analytic method would yield results similar to those of order analysis at the $p \geq .50$ ($z \geq 0.01$) level. By raising the z level, order analysis should ideally return a zero-filled matrix of order loadings.

Five factor analytic models were compared with order analysis at two levels. The frame of reference for this validation study was adapted from Armstrong and Soelberg's (1968) study. Industrial ratings of a supervisor on 20 traits by 50 employees were simulated by a table of random numbers (Rand Corporation, 1955). Adjectives used as trait descriptors were also culled from this study (p. 362) and prior to the analysis were assigned in random order to columns of the data matrix.

The only alteration of Armstrong and Soelberg's design was in dichotomization of the data matrix by an odd-even:one-zero rule.

The data were factor analyzed by principal components (PC), principal factors (PF), alpha, canonical (RAO), and image models. Order analysis was conducted at the 1.64 and 0.01 z levels. The factor analyses were based on phi correlations among the 20 "variables." With the exception of the principal components method, squared multiple correlations were used as initial communality estimates; factor analytic solutions were iterated. Retained factors were rotated by Varimax with Kaiser's normalization for all methods including order analysis.

Eigenvalues obtained for all factor analytic solutions and their order analytic counterparts (extraction indexes) are reported in Table 7.5. According to Kaiser's conservative factor extraction rule, only eigenvalues greater than 1.0 were retained. Only those latent subgroupings containing more than 5% of the total number of entities or attributes analyzed were retained; extraction indexes smaller than .05 were rejected. This decision rule is arbitrary and is modelled after the customary significance levels for rejection of the null hypothesis (Fisher, 1925).

Table 7.5 shows that by this criterion the factor analytic

Table 7.5. Eigenvalues for factor analytic solutions and their order analytic counterparts.

FACTORS	PC	PF	ALPHA	RAO	IMAGE	ORDER 0.01z	ORDER 1.64z
1	2.39	1.89	4.09	5.15	4.21	.20	.06
2	1.98	1.42	3.30	4.31	3.07	.14	**
3	1.80	1.31	2.94	3.66	2.99	.14	**
4	1.74	1.22	2.52	3.46	2.78	.14	**
5	1.48	*	2.18	2.78	2.16	.12	**
6	1.40	*	2.03	2.30	1.99	.10	**
7	1.25	*	1.67	2.24	1.92	.08	**
8	1.15	*	1.34	2.03	1.70	.05	**
9	*	*	*	*	1.41	**	**
10	*	*	*	*	1.26	**	**
11	*	*	*	*	1.20	**	**
12	*	*	*	*	1.11	**	**
13	*	*	*	*	1.06	**	**
14.	*	*	*	*	*	**	**

* Eigenvalues < 1.00
** Order analysis extraction indices < .05

methods retained from 4 to 13 factors. Principal components, alpha, and canonical factor analyses all yielded 8 factors, as did order analysis at $z \geq 0.01$. When order analysis used a $z \geq 1.64$ (.05 significance level), only one marginal dimension was identified. The percentage of total variance analyzed was greater than 66% for all factor analytic models; order analysis analyzed 1335 bits at the 0.01 z level and 5 bits at the 1.64 z level. The data matrix contained 4713 bits of variance among its attributes. Order analysis therefore analyzed 28% of the variance at the 0.01 z level and 10% at the 1.64 z level.

To avoid the necessity of presenting five separate and highly similar factor loading matrices, eight factors were rotated for the factor analytic methods and interpreted with the order analysis (1.64 z level) results. To facilitate comparison with Armstrong and Soelberg's (1968) results (derived from the same tables of random numbers as were used in the present study but obviously not using the same numbers), nine factors were rotated and interpreted for order analysis at 0.01 z level. For the same reason, the order loadings originally isolated in binary digits were divided by their row sums and square roots of these proportions were computed which further simulated factor loadings.

Table 7.6 shows a comparison of order analysis at $z \geq 0.01$

with Armstrong and Soelberg's (1968, Table 1) results. The data show that when order analysis is implemented in simulated traditional factor analytic fashion (i.e., not eliminating dimensions on the basis of statistical probability) it yields results similar to those of factor analysis. Under these circumstances, both order analysis and principal components analysis yielded "interpretable" results from randomly generated data.

Both structures in Table 7.6 are roughly comparable to the structures isolated by the other factor analytic methods reported in Table 7.7. However, the structure isolated by order analysis at 1.64 z level is different. With the exception of one variable, order analysis at the 1.64 z level (last column of Table 7.7) returned a matrix of zero loadings, which describes the generalized latent structure matrix of random data at the five per cent significance level for a one-tailed test. Thus, the probabilistic implementation of the order analysis method indicates that there are no latent dimensions of any consequence in the random data matrix.

Results anticipated on the basis of our "width of the confidence band" hypothesis were supported by this random data type of validation study. The only deviation from predicted results was observed in the case of order analysis at 1.64 z level with the variable "sensitive" which returned a value of

Table 7.6. Structures of randomly generated data matrices as analyzed by order analysis at 0.01 z level and as reported by Armstrong and Soelberg (1968, p. 362, Table 1). Order loadings were converted in this table into simulated factor loadings.

Principal Components Analysis (Armstrong and Soelberg)			Order Analysis (0.01 z level)		
I.	Fascism		I.	Democracy	
	Sensitive	-.72		Democratic	.47
	Democratic	-.72			
	Responsible	.69			
II.	Sincerity		II.	Ascendancy	
	Sincere	-.85		Aggressive	.54
				Responsive	.70
III.	Social Distance		III.	Social Distance	
	Formal	.84		Sensitive	.59
	Fair	-.65		Formal	.63
	Humble	-.54		Active	.46
IV.	Reliability		IV.	Reliability	
	Trustworthy	.80		Trustworthy	.57
	Humble	.55		Humble	.58
V.	Docility		V.	Docility	
	Patient	.69		Warm	.50
	Shallow	.59		Docile	.45
	Aggressive	-.53		Submissive	.54
VI.	Kindness		VI.	Kindness	
	Kind	.82		Patient	.47
				Fair	.58
VII.	Humor		VII.	Humor	
	Humorous	-.88		Humorous	.40
	Aggressive	-.54		Kind	.56
VIII.	Tactfulness		VIII.	Shallowness	
	Tactful	.81		Shallow	.46
IX.	Social Leadership		IX.	Social Leadership	
	Warm	.72		Sincere	.60
	Active	.71		Enthusiastic	.50
	Enthusiastic	.64		Tactful	.59
	Ambitious	.55		Ambitious	.45

Table 7.7. Structure of randomly generated data matrix as described by principal components, principal factors, alpha, canonical, and image factor analytic models and corresponding loadings as returned by order analysis at 1.64 z level. Order loadings are reported in this table as information theory bits.

Factors and High Loading Variables	Method					
	PC	PF	ALPHA	RAO	IMAGE	ORDER
I. Achievement via confirmance						
Docile	.84	.70	.74	.72	.42	.00
Ambitious	.53	.32	.31	.32	.20	.00
II. Ascendancy						
Sincere	.78	-.60	.63	-.59	.43	.00
Submissive	-.61	.55	-.53	.57	-.27	.00
III. Social distance						
Formal	.71	.55	.57	.53	.44	.00
Humble	-.69	-.58	-.57	-.57	-.44	.00
IV. Dependency						
Active	-.80	-.74	-.79	-.75	-.52	.00
Warm	.70	.59	.55	.60	.43	.00
Shallow	.50	.37	.41	.34	.36	.00
Trustworthy	.45	.37	.37	.35	.32	.00
V. Partiality						
Sensitive	.74	.54	.53	.56	.40	5.00
Fair	-.70	-.55	-.54	-.53	-.28	.00
VI. Kindness						
Kind	.74	.43	.45	.43	.39	.00
Patient	.60	.52	.58	.50	.34	.00
Responsible	.41	.33	.27	.36	.37	.00
VII. Sarcasm						
Humorous	.76	.63	.64	.63	.47	.00
Democratic	-.60	-.45	-.46	-.44	-.37	.00
Enthusiastic	-.53	-.47	-.45	-.47	-.37	.00
VIII. Tactfulness						
Aggressive	-.83	-.77	-.77	-.77	-.61	.00
Tactful	.75	.66	.63	.66	.58	.00

five bits.

It is interesting to follow the genesis of these five stray binary digits. As described by equations 6.4 and 6.16, we are attempting to generalize isolated manifest and latent structures of entities and attributes by using multiple z tests, a practice usually rejected at the elementary statistics course level. Generalization of the square dominance matrix M_{rxr} (in this case of the order 50), necessitates $50^2 =$ 2500 comparisons of its elements, i.e., 2500 multiple z tests. As can be expected, five per cent of these tests will be randomly significant, which leaves 125 error bits added to the total amount of binary variance contained by the generalized dominance matrix for entities ${}^{p}P_{rxr}$ if the redundance-removal transformation (as in equation 6.18) is used. Five per cent of these randomly generated binary digits (6.25 bits) should reappear in the matrices representing the generalized relational space of attributes ${}^{d_k}G_{qxq}$ (equation 6.17, transformation 6.18) and, consequently, in the resultant F_{qxd} matrix of order loadings. The present analysis actually found 135 randomly generated binary digits contained in matrix ${}^{p}S_{rxr}$ and 5 random bits contained in the F_{qxd} matrix.

The implementation of multiple z tests in the present model of order analysis was dictated by its simplicity with respect to its actual computer program realization; any attempt for addition of some type of ANOVA contrast seems premature at the present state of development. However desirable this implementation seems to be, its realization would be mostly of theoretical significance. Dominance matrices for actual data ordinarily contain several thousand bits of information and the amount of error variation in terminal matrices, generated by implementation of multiple z type generalization procedures, is negligible. However, when the isolation of manifest structures is attempted, the amount of error variation generated by the multiple z tests should be taken into consideration.

Chapter 8

RETROSPECT AND PROSPECTUS

Standardization

Despite the singleminded effort for nearly half a decade, the development of order analysis as a method for a general, semantically based account of dimensional arrangement of cognitive fields, is far from completion. The initial adoption of the relational matrix of the negative implication function and its converse influenced the formulation of the deterministic model of order analysis to such a degree, that the early computerized version of order analysis was referred to as Negative Implication Analysis (Krus, 1972). This does not mean that this particular operation cannot be altered.

When the data matrix is composed of roughly comparable entities and attributes, the preservation of its elevation and scatter (Cronbach & Gleser, 1953) by the dominance matrix can be an advantage in certain situations. As contrasted with standardized correlation matrices of factor analysis, this preservation of elevation and scatter can be incorrect when the data is presented in markedly incomparable units. Standardization of the data matrix prior to analysis is also recommended for raw score factor analysis (Nunnally, 1967). It is interesting to observe that the dichotomization procedure approximates direct standardization of the data matrix. In the case of markedly disparate data, the dichotomization according to median, or a double standardization of the data matrix is recommended.

Optimization of the dominance-consonance ratios.

Another point open to discussion is the inability of the dominance type of relational matrix to optimize the dominance-consonance ratios. This can be understood by considering primitive bases of relational matrices used in various methods of dimensional analyses.

Eventually, every relational matrix can be decomposed into a combination of some of the sixteen basic functions of the logical calculus. These logical functions can be classified into order-independent and order-dependent aliorelative classes. Coombs' classification of relations into the dominance (order) and consonance (proximity) classes reflects these underlying logical classifications. (Coombs, Dawes, and Tversky, 1970, p. 33).

For example, a nucleus of the phi coefficient of a fourfold point surface (Yule, 1919, pp. 60-75) is based on the value of the determinant of this square matrix (cf. Lazarsfeld, 1961, pp. 111-157). Close scrutiny of the determinant of this type of matrix will reveal that frequencies of equivalence (i.e., one-one and zero-zero tuples) are compared with frequencies of negative implication and its converse, counted across one-zero and zero-one tuples. Preponderance of either equivalence or negative implication binary relations will determine whether the matrix of phi coefficients will be of dominance or consonance type with respect to relations between the elements of the original data matrix. In a similar fashion, a reflection of a variable in a data matrix changes its relation with any other variable from dominance to consonance or vice versa. As contrasted with this procedure, the dominance matrix does not weigh these ratios. This is e.g. reflected by absence of negative order loadings in the F_{qxd} or F_{rxd} matrices (negative order loadings, as reported elsewhere, were created by rotational procedures). On the other hand, the dominance matrix preserves the ordinal properties of the data matrix better than a symmetric matrix of correlations. Thus any future attempt to optimize the data matrix orientation and preserve at the same time its potential orders, would lead to a relational matrix superior to either dominance or correlation-type relational matrices.

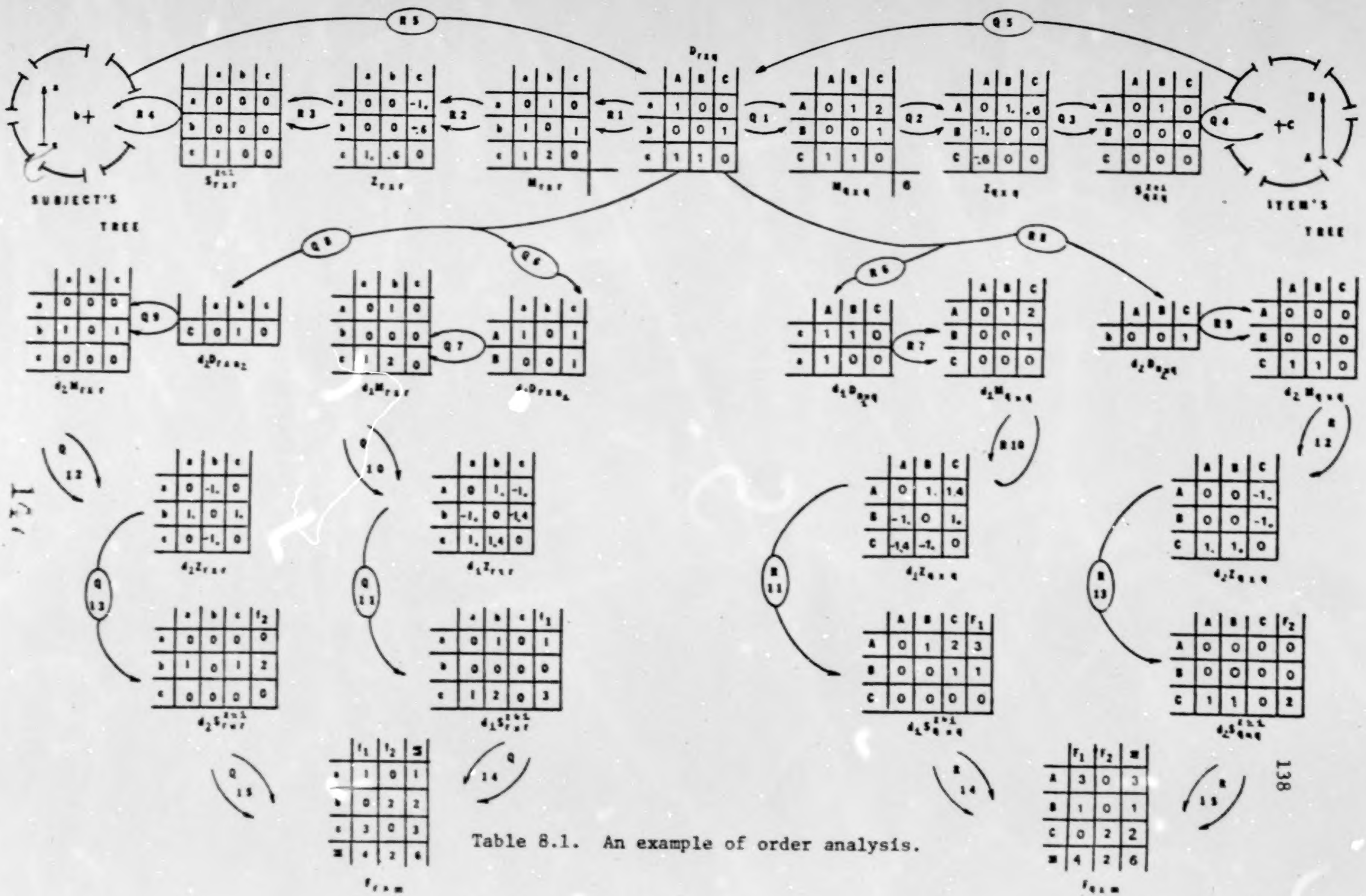
Thurstone's Law of Comparative Judgement.

In its least general form (case 5), Thurstone's Law of Comparative Judgment (Thurstone, 1927, a, b, c) states that the difference between the scale values of two stimuli ($S_j - S_i$) is a function of the proportion of times stimulus j is judged greater than stimulus i. The Law of Comparative Judgment represents the first systematic use of the difference relation properties in the field of psychological measurement; it is the direct antecedent of the concept of dominance matrices, as used by Coombs (1964) in his data theory and by order analysis.

Types of variance extracted.

It is necessary to explicate the types of variation which the present model partitions. It should be remembered that the variance is defined here in terms of directional one-zero changes, i.e. as a number of bits in the test space or its part. Variance conventionally expressed as a second moment about the mean is a function of this information measure and has a different metric. The advantage of order analysis of dominance matrices is that it permits a precise, undistorted manipulation of basic relations between the elements of the test space, with the open option for rescaling prior to or after the analysis proper.

To facilitate our discussion, let us consider the contrived example in Table 8.1. The data matrix in this example D_{rxq} has its total variance equal to twenty bits with row and column variances equal to six bits each, which leaves eight bits for interaction. The row and column variance is recorded in M_{rxr} and M_{qxq} matrices and the interaction variance could be described in terms of corresponding minor matrices of magnitudes.


R 5
Q 5
SUBJECT'S TREE
ITEM'S TREE
a
b+
c
+C
B
A
R 4
R 3
R 2
R 1
Q 1
Q 2
Q 3
Q 4
G
Q 8
Q 6
Q 9
Q 7
R 8
R 6
R 7
R 9
Q 12
Q 10
Q 13
Q 11
Q 15
Q 14
R 10
R 12
R 11
R 13
R 14
R 15

$S^{z\geq1}_{r\times r}$

	a	b	c
a	0	0	0
b	0	0	0
c	1	0	0

$Z_{r\times r}$

	a	b	c
a	0	0	-1.
b	0	0	-.6
c	1.	.6	0

$M_{r\times r}$

	a	b	c
a	0	1	0
b	1	0	1
c	1	2	0

$D_{r\times q}$

	A	B	C
a	1	0	0
b	0	0	1
c	1	1	0

$M_{q\times q}$

	A	B	C
A	0	1	2
B	0	0	1
C	1	1	0

$Z_{q\times q}$

	A	B	C
A	0	1.	.6
B	-1.	0	0
C	-.6	0	0

$S^{z\geq1}_{q\times q}$

	A	B	C
A	0	1	0
B	0	0	0
C	0	0	0

$d_2M_{r\times r}$

	a	b	c
a	0	0	0
b	1	0	1
c	0	0	0

$d_2D_{r\times a_2}$

	a	b	c
C	0	1	0

$d_1M_{r\times r}$

	a	b	c
a	0	1	0
b	0	0	0
c	1	2	0

$d_1D_{r\times a_1}$

	a	b	c
A	1	0	1
B	0	0	1

$d_1D_{a_1\times q}$

	A	B	C
c	1	1	0
a	1	0	0

$d_1M_{q\times q}$

	A	B	C
A	0	1	2
B	0	0	1
C	0	0	0

$d_2D_{a_2\times q}$

	A	B	C
b	0	0	1

$d_2M_{q\times q}$

	A	B	C
A	0	0	0
B	0	0	0
C	1	1	0

$d_2Z_{r\times r}$

	a	b	c
a	0	-1.	0
b	1.	0	1.
c	0	-1.	0

$d_1Z_{r\times r}$

	a	b	c
a	0	1.	-1.
b	-1.	0	-1.4
c	1.	1.4	0

$d_1Z_{q\times q}$

	A	B	C
A	0	1.	1.4
B	-1.	0	1.
C	-1.4	-1.	0

$d_2Z_{q\times q}$

	A	B	C
A	0	0	-1.
B	0	0	-1.
C	1.	1.	0

$d_2S^{z\geq1}_{r\times r}$

	a	b	c	f_2
a	0	0	0	0
b	1	0	1	2
c	0	0	0	0

$d_1S^{z\geq1}_{r\times r}$

	a	b	c	f_1
a	0	1	0	1
b	0	0	0	0
c	1	2	0	3

$d_1S^{z\geq1}_{q\times q}$

	A	B	C	F_1
A	0	1	2	3
B	0	0	1	1
C	0	0	0	0

$d_2S^{z\geq1}_{q\times q}$

	A	B	C	F_2
A	0	0	0	0
B	0	0	0	0
C	1	1	0	2

$F_{r\times m}$

	f_1	f_2	Σ
a	1	0	1
b	0	2	2
c	3	0	3
Σ	4	2	6

$F_{q\times m}$

	F_1	F_2	Σ
A	3	0	3
B	1	0	1
C	0	2	2
Σ	4	2	6

Table 8.1. An example of order analysis.

If we partitioned the data matrix into the vector supermatrix (a procedure analogous to the principal components analysis), the minor matrices would change into the zero matrices, the total variance would be equal in every type of analysis to its corresponding matrix of magnitudes and the number of dimensions to r or q. Observe the R chain in Table 8.1 through steps 1 to 8, which partitions the data matrix D_{rxq} into a supermatrix with ${}_{d_1}D_{n1xq}$ and ${}_{d_2}D_{n2xq}$ submatrices. The total variance of the ${}_{d_1}D_{n1xq}$ submatrix is 9 and consists of row variance of four bits (${}_{d_1}M_{qxq}$), column variance of one bit ($S^{z\ 1}_{rxr}$) and interaction variance of four bits. The total variance of ${}_{d_2}D_{n2xq}$ subvector is 2. All variance in this subvector is due to row variation only. The total variance of the whole supermatrix is therefore eleven bits, composed of the six bits of row variation, one bit of column variation and six bits of interaction. Our example was purposefully constructed in such a way, that the same variance distribution can be found in the Q branch of our example. This is of course not necessarily the case. The sum of total variances of both R and Q branches is therefore twenty-two bits, two bits more than the total variance of the original data matrix. This inequality can be corrected by subtracting one bit for mutually shared $S^{z\geq 1}_{rxr}$ and $S^{z\geq 1}_{qxq}$ variance; which was introduced by concatenation of Guttman-type data vectors. (Two vectors were catenated in our example in every branch and variation created by this chaining stands for jumps typical of this data type (one-zero change as circled below:

...111110000...

...111100000...).

The probabilistic character of this mutually shared variance at the first level of dimensionality estimation is due to its prior Z_{qxq} or Z_{rxr} transformation. Let us restrict our scrutiny to R-type covariation design. The initial partitioning of the data matrix, derived from the structural characteristics of the $S^{z\geq1}_{rxr}$ matrix is based on the generalization of the structural properties of the M_{rxr} matrix to the structure of a population of subjects by the Z_{rxr} transformation. The partitioned variance is interdimensional. Successive steps (R10 through R13, Table 8.1) separately repeat this estimation process for every $d_1S^{z\geq1}_{qxq}$, $d_2S^{z\geq1}_{qxq}$, ..., $d_kS^{z\geq1}_{qxq}$, ..., $d_mS^{z\geq1}_{qxq}$ matrix. These matrices are the generalizations of the structural properties of the corresponding M_{qxq} matrices to the structure of a population of variables by the d_1Z_{qxq}, d_2Z_{qxq}, ..., d_kZ_{qxq}, ..., d_mZ_{qxq} transformations. The partitioned variance is intradimensional.

Integrated analysis of the general relational space.

Another interesting point is the affinity of order analysis to the logic behind Campbell and Fiske's (1959) multitrait-multimethod matrix, which is the specific formulation of the general relational space of attributes and entities. Consider matrix elements a, b, c, ..., i of the square matrix of the order three. The complete relational space of this matrix is depicted in Table 8.2. The R and C symbols stand for the relations between its respective row and column elements. Its heterotrait-monomethod triangles are equivalent to the upper and lower triangular matrices of our M_{rxr} matrix and its validity diagonal contains the values of M_{qxq} matrix. The requirements for convergent and discriminant validity could be phrased in terms of factorial validity as appropriate inter- and intra-dimensional generalizability.

The iterative and integrated analysis of this type of space could be imagined as simultaneous analysis in R and Q branches (Table 8.1), with the terminal matrix of latent attributes and entities defined in three dimensions as F_{rxqxd}. Iteration criterion could be the maximization of variance for all four structures (manifest attributes, manifest entities, latent attributes, latent entities) and minimization of the interaction variance (blank spaces in Table 8.2).

Considering the development in order analysis from its deterministic model (Krus and Bart, 1972) through its probabilistic formulation (Krus, 1973) and technical development (Krus, 1973), subsequent elaboration will probably entail some improve-

Table 8.2. Relational space of a square matrix, depicting the relations between its elements. R subscripted symbols stand for successive row elements, C for column elements. There are three rows and three columns with elements a, b, c, ..., i - elements being labeled by starting in uppermost left corner and by continuing rowwise. Blank spaces stand for interaction relations and V symbols for homorelative relations.

	a	b	c	d	e	f	g	h	i
a	V	R_1	R_1	C_1			C_1		
b	R_1	V	R_1		C_2			C_2	
c	R_1	R_1	V			C_3			C_3
d	C_1			V	R_2	R_2	C_1		
e		C_2		R_2	V	R_2		C_2	
f			C_3	R_2	R_2	V			C_3
g	C_1			C_1			V	R_3	R_3
h		C_2			C_2		R_3	V	R_3
i			C_3			C_3	R_3	R_3	V

ments suggested here as well as reconsideration of the challenging issues connected with the general problem of structural analysis and generalization; questions of optimization versus maximization of root extraction procedures and redundancy removal transformations.

The most important property of order analysis is perhaps its insensitivity to random variation patterns as compared with other models of multivariate analysis. This property is supported by our theoretical reasoning and experimental validation and has profound implications for the general field of theoretical and applied multivariate measurement.

References

Adams, E., Fagot, R. F., Robinson, R. E. A theory of appropriate statistics. *Psychometrika*, 1965, *30*, 99-128.

Airasian, P. W., Bart, W. M. Ordering theory: a new and useful measurement model. *Educational Technology*, 1973, *5*, 56-60.

Aristotle. *Selected works*. Ed. by W. D. Ross. New York: Scribner, 1927.

Armstrong, J. W. & Soelberg, P. On the interpretation of factor analysis. *Psychological Bulletin,* 1968, *70*, 361-364.

Bart, W. M. & Krus, D. J. An ordering-theoretic method to determine hierarchies among items. *Educational and Psychological Measurement*, 1973, *33*, 291-300.

Bartlett, M. W. Tests of significance in factor analysis. *British Journal of Psychology*, Statistical Section, 1950, *3*, 77-85.

Beals, R., Krantz, D. H., & Tversky, A. Foundations of multidimensional scaling. *Psychological Review*, 1968, *75*, 125-142.

Bennett, J. F. A method for determining the dimensionality of a set of rank orders. Unpublished doctoral dissertation, University of Michigan, 1951.

Bennett, J. F. Determination of the number of independent parameters of a score matrix from the examination of rank orders. *Psychometrika,* 1956, *21*, 383-395.

Boole, G. *An investigation of the laws of thought*. (1854). New York, Dover, 1959.

Borgen, F. H. & Weiss, D. J. Cluster analysis and counseling research. *Journal of Counseling Psychology*, 1971, *18*, 583-591.

Campbell, D. T., Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*, 81-105.

Campbell, N. R. *An account of the principles of measurement and calculation.* London: Longmans Green, 1928.

Carnap, R. Logical foundations of probability. Chicago: University of Chicago Press, 1950.

Cattell, R. B. The sixteen personality questionnaire. Champaign, Ill.: Institute for Personality and Ability Testing, 1957.

Cattell, R. B. Description and measurement of personality (1946). New York: Johnson Reprint Corporation, 1969.

Cayley, A. On the analytical forms called trees, with applications to the theory of chemical combinations. Reports of the British Association for the Advancement of Science, 1875, 257-305.

Cliff, N. Scaling. In P. H. Mussen and M. R. Rozenzweig (Eds.) Annual Review of Psychology, 1973, 24, 473-506.

Cohen, M. R., Nagel, G. An introduction to logic. (1934). New York: Harcourt, 1962.

Coombs, C. H. A theory of data. New York: Wiley, 1964.

Coombs, C. H., Dawes, R. M., & Tversky, A. Mathematical psychology. Englewood Cliffs, N. J.: Prentice Hall, 1970.

Cournot, A. A. Exposition de la theorie des chances. Paris, 1843.

Cronbach, L. J., Gleser, G. C. Assessing similarity between profiles. Psychological Bulletin, 1953, 50, 456-473.

Cureton, E. E., Mulaik, S. A. On simple structure and the solution to Thurstone's "Invariant" Box Problem. Multivariate Behavioral Research, 1971, 6, 375-387.

DeBruihn, N. G., Morselt, B. J. M. A note on plane trees. Journal of Combinatorial Theory, 1967, 2, 27-34.

Edwards, A. L. Techniques of attitude scale construction. New York: Appleton-Century-Crofts, 1957.

Euclid. Elements. Ed. by Thomas L. Heath. New York: Dover, 1926.

Festinger, Leon. The treatment of qualitative data by "scale analysis." Psychological Bulletin, 1947, 44, 149-161.

Fechner, G. T. Zur experimentalen Aesthetik. *Abhandlung fur mathematische physik.* 1871, *9*, 555-635.

Fisher, R. A. *Statistical methods for research workers*. London: Oliver and Boyd, 1925.

Frege, G. *Grundlagen der Arithmetik, eine logish-mathematische Untersuchung uber den Begriff der Zahl* (1884). Hildesheim: Georg Olms Verlagsbuchhandlung, 1961.

Frege, G. *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet* (1893). Hildesheim: Georg Olms Verlagsbuchhandlung, 1962.

Goldberg, L. R. A model of item ambuity in personality assessment. *Educational and Psychological Measurement*, 1963, *23*, 467-492.

Gibson, W. A. Nonlinear factors in two dimensions. *Psychometrika*, 1960, *25*, 381-392.

Gibson, W. A. Three multivariate models: Factor analysis, Latent structure analysis, and Latent Profile Analysis. In P. F. Lazarsfeld, Ed. *Readings in Mathematical Social Science*. Cambridge, Mass.: The M.I.T. Press, 1966.

Goodenough, W. H. A technique for scale analysis. *Educational and Psychological Measurement*, 1944, *4*, 179-190.

Guilford, J. P. *Psychometric Methods*. New York: McGraw, 1936.

Guilford, J. P. & Martin, H. G. *The Guilford-Martin inventory of factors GAMIN: Manual of directions and norms*. Beverly Hills, California: Sheridan Supply Co., 1943.

Guilford, J. P. *Psychometric methods*. (Second edition). New York: McGraw, 1954.

Guilford, P. J. When not to factor analyze. In D. N. Jackson and S. Messick (Ed.) *Problems in human assessment*. New York: McGraw, 1967.

Gulliksen, H. *Theory of mental tests*. New York: Wiley, 1950.

Guttman, L. The quantification of a class of attributes: A theory and method of scale construction. In P. Horst (Ed.) *The prediction of personal adjustment.* New York: Social Science Research Council, 1941.

Guttman, L. A revision of Chapin's social status scale. *American Sociological Review*, 1942, *7*, 362-369.

Guttman, L. A. A basis for scaling qualitative data. *American Sociological Review*, 1944, *9*, 139-150.

Guttman, L. An approach for quantifying paired comparisons and rank order. *Annals of the Mathematical Statistics*, 1946, *17*, 144-163.

Guttman, L. The Cornell technique for scale and intensity analysis. *Educational and Psychological Measurement*, 1947, *7*, 247-280.

Guttman, L. Chapters in S. A. Stoufer (Ed.) *Measurement and prediction*. New York: Wiley, 1950.

Guttman, L. The principal components of scalable attitudes. In P. F. Lazarsfeld (Ed.) *Mathematical thinking in the social science.* New York: Russell, 1954.

Guttman, L. A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, 1968, *33*, 469-506.

Hall, C. S., Lindzey, G. *Theories of personality*. New York: Wiley, 1957.

Harman, H. H. *Modern factor analysis*. Chicago: The University of Chicago Press, 1960.

Harris, C- W. Some Rao-Guttman relationships. *Psychometrika*, 1962, *27*, 247-263.

Hempel, C. G. *Fundamentals of concept formation in empirical science*. International Encyclopedia of United Science. Chicago: University of Chicago Press, 1952.

Horst, P. *Factor analysis of data matrices*. New York: Holt, 1965.

Horst, P. An overview of the essentials of multivariate analysis methods. In R. B. Cattell (Ed.) *Handbook of multivariate experimental psychology.* Chicago: Rand McNally, 1966.

Humphreys, L. G., Ilgen, D. R. Note on a criterion for the number of common factors. *Educational and Psychological Measurement*, 1969, *29*, 571-578.

Huntington, E. U. *The continuum and other types of serial order.* Cambridge, Mass.: Harvard University Press, 1917.

Jaeger, W. *Aristoteles.* Berlin: Springer Verlag, 1923.

Kant, E. *Selected works*. Ed. by T. M. Greene. New York: Scribner, 1929.

Kaiser, H. F. Psychometric approaches to factor analysis. In Anastasi, A. (Ed.). *Testing Problems in Perspective.* Washington: American Council on Education, 1966.

Kemeny, J. G., Schleifer, A., Snell, J. L., Thompson, G. L. *Finite mathematics.* Englewood Cliffs, N. J.: Prentice Hall, 1972.

Krantz, D. H., Luce, D. R., Suppes, P., Tversky, A. *Foundations of measurement*. New York: Academic Press, 1971.

Krus, D. J. Negative Implication Analysis: A model for multidimensional analysis and scaling of dichotomous data. Fortran program for CDC 6600 computer. Version 1.1. *University of Minnesota Computer Center Tape Library,* SN2776, 1972.

Krus, D. J. Order analysis: A Fortran program for generalizable multidimensional analysis of binary data matrices. Minneapolis: University of Minnesota, Research and Development Center, *Technical Report*, OP-22, 1973.

Krus, D. J., Bart, W. M. An ordering-theoretic method of multidimensional scaling of items. Minneapolis: University of Minnesota, Research and Development Center, *Technical Report*, RR-33, 1972.

Kruskal, J. B. Nonmetric multidimensional scaling. *Psychometrika*, 1964, *29*, 115-130.

Lazarsfeld, P. F. Latent structure analysis. In S. Koch (Ed.) *Psychology: A study of a science*, Vol. III, New York: McGraw, 1958.

Lazarsfeld, P. F. The algebra of dichotomous systems. In H. Solomon (Ed.) *Studies in item analysis and prediction.* Stanford, Calif.: Stanford University Press, 1961.

Leibnitz, G. W. *Selected works*. Ed. by P. P. Wiener, New York, Scribner, 1951.

Lingoes, J. C. An IBM program for Guttman-Lingoes smallest space analysis. Behavioral Science, 1966, 11, 33.

Loevinger, J. The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. Psychological Bulletin, 1948, 45, 507-530.

Lord, F. A theory of test scores. Psychometric Monographs, 7, 1952.

Lord, F. The relation of test score to the trait underlying the test. In P. F. Lazarsfeld, Ed. Readings in mathematical social science. Cambridge, Mass.: The M.I.T. Press, 1966.

Lord, F. M., Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

Luce, R. D. Tukey, J. W. Simultaneous conjoint measurement: a new type of fundamental measurement. Journal of Mathematical Psychology, 1964, 1, 1-27.

Lunneborg, C. E. Dimensional analysis, latent structure and the problem of patterns. Dissertation Abstracts, 1960, 20, 4186.

Lykken, D. T. A method of actuarial pattern analysis. Psychological Bulletin, 1956, 53, 102-107.

McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika, 1947, 12, 153-157.

McNemar, Q. Psychological statistics. New York: Wiley, 1969.

McQuitty, L. L. A method of pattern analysis for isolating typological and dimensional constructs. San Antonio, Texas: Air Force Headquarters, Personnel and Training Research Center, Research Bulletin, TN-62, 1955.

McQuitty, L. L. Isolating predictor patterns associated with major criterion patterns. Educational and Psychological Measurement, 1957, 17, 3-42.

Meehl, P. E. Configural scoring. Journal of Consulting Psychology, 1950, 14, 165-171.

Mendelson, E. Boolean algebras and switching circuits. New York: McGraw, 1970.

Milholland, J. E. Dimensionality of response patterns for the method of single stimuli. Unpublished doctoral dissertation, University of Michigan, 1953.

Nie, N. H., Bent, D. H., & Hull, C. H. *Statistical package for the social sciences.* New York: McGraw, 1970.

Nunnally, J. *Psychometric Theory*. New York: McGraw, 1967.

Pemberton, C. The closure factors related to temperament, *Journal of Personality*, 1952, *21*, 159-175.

Rand Corporation. *A million random digits with 100,000 normal deviates.* New York: Free Press, 1955.

Rao, C. R. Estimation and tests of significance in factor analysis. *Psychometrika,* 1955, *20*, 93-111.

Rasch, G. *Probabilistic models for some intelligence and attainment tests.* Copenhagen: Danish Institute for Educational Research, 1960.

Read, R. C. The coding of various kinds of unlabeled trees. In R. C. Read (Ed.) *Graph theory and computing.* New York: Academic Press, 1972.

Rummel, R. J. *Applied factor analysis.* Evanston: Northwestern University Press, 1970.

Russell, B. *Principles of mathematics*. (1903). New York: Norton, 1938.

Russell, B. *Introduction to Mathematical Psychology.* (1919). New York: Simon and Schuster, 1971.

Shannon, C. E. & Weaver, W. *The mathematical theory of communication.* Urbana, Ill.: The University of Illinois Press, 1949.

Shepard, R. N. The analysis of proximities. *Psychometrika,* 1962, *27*, 125-140.

Spearman, C. General intelligence objectively determined and measured. *American Journal of Psychology*, 15, 201-292, 1904.

Spearman, C. *The abilities of man.* (1926). New York: Macmillan, 1927.

Stevens, S. S. Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.) *Handbook of experimental psychology.* New York: Wiley, 1951.

Sullivan, H. S. *The interpersonal theory of psychiatry.* New York: Norton, 1953.

Suppes, P., Zinnes, J. L. Basic measurement theory. In R. D. Luce, R. R. Busch, E. Galanter (Eds.) *Handbook of mathematical psychology.* New York: Wiley, 1963.

Thalwitzer, K. Characterizierung eines Labyrinthes durch ein Wort uber einem zweistellige Alphabet. In H. Sachs, H. J. Voss, H. Walther (Eds.) *Beitrage zur Graphentheorie.* Leipzig: Teubner, 1968.

Thorndike, R. M. Method of extraction, type of data, and adequacy of solutions in factor analysis. University of Minnesota, Unpublished Doctoral Dissertation, 1970.

Torgerson, W. S. *Theory and methods of scaling.* New York: Wiley, 1958.

Torgerson, W. S. Multidimensional scaling of similarity. *Psychometrika*, 1965, *30*, 379-393.

Thurstone, L. L. The method of paired comparisons for social values. *Journal of Abnormal and Social Psychology*, 1927a, *4*, 384-400.

Thurstone, L. L. Equally often noticed differences. *Journal of Educational Psychology*, 1927b, *18*, 289-293.

Thurstone, L. L. A law of comparative judgment. *Psychological Review*. 1927c, *34*, 273-286.

Thurstone, L. L. *Multiple factor analysis.* Chicago: The University of Chicago Press, 1947.

Weiss, D. J. Factor analysis and counseling research. *Journal of Counseling Psychology*, 1970, *17*, 5, 477-485.

Weiss, D. J. Further considerations in applications of factor analysis. *Journal of Counseling Psychology*, 1971, *18*, 85-92.

White, B. W., Saltz, E. The measurement of reproducibility. *Psychological Bulletin*, 1957.

Whitehead, A. N. Introduction to Mathematics. (1911) New York: Oxford University Press, 1969.

Wiggins, J. S. Personality structure. In P. R. Farnsworth, M. R. Rozenzweig, J. T. Polefka (Eds.) Annual Review of Psychology. Palo Alto: Annual Reviews, 1968.

Yule, G. U. Introduction to the theory of statistics. London: Charles Griffin, 1919.

31. J. Turnure & M. Thurlow. *The effects of interrogative elaborations on the learning of normal and EMR children.* Research Report #44. January 1973. (Proceedings of the International Association for the Scientific Study of Mental Deficiency, in press).

32. J. Turnure & S. Samuels. *Attention and reading achievement in first grade boys and girls.* Research Report #43. November 1972. (*Journal of Educational Psychology*, in press).

33. R. Riegel, A. Taylor, S. Clarren, & F. Danner. *Training educationally handicapped children to use associative grouping strategies for the organization and recall of categorizable materials.* Research Report #42. November 1972.

34. R. Riegel, F. Danner, & A. Taylor. *Steps in sequence: Training educationally handicapped children to use strategies for learning.* Development Report #2. November 1972.

35. A. Taylor, M. Thurlow, & J. Turnure. *The teacher's introduc[illegible] The math vocabulary program.* Development Report #1. March 1973.

36. J. Turnure & M. Thurlow. *The effects of structural variations in elaboration on learning by normal and EMR children.* Research Report #41. September 1972.

37. A. Taylor & N. Bender. *Variations of strategy training and the recognition memory of EMR children.* Research Report #40. September 1972. (American Educational Research Journal, in press).

38. D. Moores, C. McIntyre, & K. Weiss. *Evaluation of programs for hearing impaired children: Report of 1971-72.* Research Report #39. September 1972.

39. R. Rubin. *Follow-up of applicants for admission to graduate programs in special education.* Occasional Paper #11. July 1972.

40. D. Moores. *Communication - Some unanswered questions and some unquestioned answers.* Occasional Paper #10. July 1972.

41. A. Taylor & S. Whitely. *Overt verbalization and the continued production of effective elaborations by EMR children.* Research Report #38. June 1972. (*American Journal of Mental Deficiency*, in press).

42. R. Riegel. *Measuring educationally handicapped children's organizational strategies by sampling overt groupings.* Research Report #37. May 1972.

43. E. Gallistel, M. Boyle, L. Curran, & M. Hawthorne. *The relation of visual and auditory aptitudes to first grade low readers' achievement under sight-word and systematic phonic instruction.* Research Report #36. May 1972.

44. E. Gallistel & P. Fischer. *Decoding skills acquired by low readers taught in regular classrooms using clinical techniques.* Research Report #35. May 1972.

45. J. Turnure & M. Thurlow. *Verbal elaboration in children: Variations in procedures and design.* Research Report #34. March 1972.

46. D. Krus & W. Bart. *An ordering-theoretic method of multidimensional scaling of items.* Research Report #33. March 1972.

47. J. Turnure & S. Larsen. *Effects of various instruction and reinforcement conditions on the learning of a three-position oddity problem by nursery school children.* Research Report #32. March 1972.

48. J. Turnure & S. Larsen. *Outerdirectedness in mentally retarded children as a function of sex of experimenter and sex of subject.* Research Report #31. March 1972.

49. J. Rynders & J. Horrobin. *A mobile unit for delivering educational services to Down's Syndrome (Mongoloid) infants.* Research Report #30. January 1972. (Presented at Council for Exceptional Children, Special National Conference, Memphis, December, 1971).

50. F. Danner & A. Taylor. *Pictures and relational imagery training in children's learning.* Research Report #29. December 1971. (*Journal of Experimental Child Psychology*, in press).

51. J. Turnure & M. Thurlow. *Verbal elaboration phenomena in nursery school children.* Research Report #28. December 1971. (Study II: *Proceedings of 81st Annual Convention of the American Psychological Association*, in press).

52. D. Moores & C. McIntyre. *Evaluation of programs for hearing impaired children: Progress report 1970-71.* Research Report #27. December 1971.

53. S. Samuels. *Success and failure in learning to read: A critique of the research.* Occasional Paper #9. November 1971. (In M. Kling, the Literature of Research in Reading with Emphasis on Modes, Rutgers University, 1971).

54. S. Samuels. *Attention and visual memory in reading acquisition.* Research Report #26. November 1971.

55. J. Turnure & M. Thurlow. *Verbal elaboration and the promotion of transfer of training in educable - mentally retarded children.* Research Report #25. November 1971. (*Journal of Experimental Child Psychology*, 1973, *15*, 137-148.

56. A. Taylor, M. Josberger, & S. Whitely. *Elaboration training and verbalization as factors facilitating retarded children's recall.* Research Report #24. October 1971. (Journal of Educational Psychology, in press).

57. W. Bart & D. Krus. *An ordering-theoretic method to determine hierarchies among items.* Research Report #23. September 1971.

58. A. Taylor, M. Josberger, & J. Knowlton. *Mental elaboration and learning in retarded children.* Research Report #22. September 1971. (Mental Elaboration and Learning in EMR children. *American Journal of Mental Deficiency*, 1972, *77*, 69-76).

59. J. Turnure & S. Larsen. Outerdirectedness in educable mentally retarded boys and girls. Research Report #21. September 1971. (American Journal of Mental Deficiency, in press).

60. R. Bruininks, T. Glaman, & C. Clark. Prevalency of learning disabilities: Findings, issues, and recommendations. Research Report #20. June 1971. (Presented at Council for Exceptional Children Convention, Miami Beach, April, 1971).

61. M. Thurlow & J. Turnure. Mental elaboration and the extension of mediational research: List length of verbal phenomena in the mentally retarded. Research Report #19. June 1971. (Journal of Experimental Child Psychology, 1972, 14, 184-195.

62. G. Siegel. Three approaches to speech retardation. Occasional Paper #8. May 1971.

63. D. Moores. An investigation of the psycholinguistic functioning of deaf adolescents. Research Report #18. May 1971. (Exceptional Children, May 1970, 36, 645-652).

64. D. Moores. Recent research on manual communication. Occasional Paper #7. April 1971. (Keynote Address, Division of Communication Disorders, Council for Exceptional Children Annual Convention, Miami Beach, April 1971).

65. J. Turnure, S. Larsen, & M. Thurlow. Two studies on verbal elaboration in special populations. I. The effects of brain injury II. Evidence of transfer of training. Research Report #17. April 1971. (Study I: American Journal of Mental Deficiency, in press).

66. R. Bruininks & J. Rynders. Alternatives to special class placement for educable mentally retarded children. Occasional Paper #6. March 1971. (Focus on Exceptional Children, 1971, 3, 1-12).

67. D. Moores. Neo-oralism and the education of the deaf in the Soviet Union. Occasional Paper #5. February 1971. (Exceptional Children, January 1972, 39, 377-384).

68. D. Feldman, B. Marrinan, & S. Hartfeldt. Unusualness, appropriateness, transformation and condensation as criteria for creativity. Research Report #16. February 1971. (American Educational Research Association Annual Conference, New York, February 1971).

69. P. Broen & G. Siegel. Variations in normal speech disfluencies. Research Report #15. January 1971. (Language & Speech, in press).

70. D. Feldman. Map understanding as a possible crystallizer of cognitive structures. Occasional Paper #4. January 1971. (American Educational Research Journal, 1971, 3, 484-502).

71. J. Rynders. Industrial arts for elementary mentally retarded children: An attempt to redefine and clarify goals. Occasional Paper #3. January 1971.

72. D. Moores. Education of the deaf in the United States. Occasional Paper #2. November 1970. (Moscow Institute of Defectology, 1971, published in Russian).

73. R. Bruininks & C. Clark. Auditory and learning in first-, third-, and fifth-grade children. Research Report #14. November 1970.

74. R. Bruininks & C. Clark. Auditory and visual learning in first grade educable mentally retarded normal children. Research Report #13. November 1970. (American Journal of Mental Deficiency, 1972, 76, No. 5, 561-567).

75. R. Bruininks. Teaching word recognition to disadvantaged boys with variations in auditory and visual perceptual abilities. Research Report #12. November 1970. (Journal of Learning Disabilities, 1970, 3, 30-39).

76. R. Bruininks & W. Lucker. Change and stability in correlations between intelligence and reading test scores among disadvantaged children. Research Report #11. October 1970. (Journal of Reading Behavior, 1970, 2, 295-305).

77. R. Rubin. Sex differences in effects of kindergarten attendance on development of school readiness and language skills. Research Report #10. October 1970. (Elementary School Journal, 72, No. 5, February 1972).

78. R. Rubin & B. Balow. Prevalence of school learning & behavior disorders in a longitudinal study population. Research Report #9. October 1970. (Exceptional Children, 1971, 38, 293-299).

79. D. Feldman & J. Bratton. On the relativity of giftedness: An empirical study. Research Report #8. August 1970. (American Educational Research Annual Conference, New York, February 1971).

80. J. Turnure, M. Thurlow, & S. Larsen. Syntactic elaboration in the learning & reversal of paired-associates by young children. Research Report #7. January 1971.

81. R. Martin & L. Berndt. The effects of time-out on stuttering in a 12-year-old boy. Research Report #6. July 1970. (Exceptional Children, 1970, 37, 303-304).

82. J. Turnure & M. Welsh. The effects of varied levels of verbal mediation on the learning and reversal of paired associates by educable mentally retarded children. Research Report #5. June 1970. (Study I: American Journal of Mental Deficiency, 1971, 76, 60-67. Study II: American Journal of Mental Deficiency, 1971, 76, 306-312).

83. J. Turnure, J. Rynders, & N. Jones. Effectiveness of manual guidance, modeling & trial & error learning for inducing instrumental behavior in institutionalized retardates. Research Report #4. June 1970. (Merrill-Palmer Quarterly, 1973, 19, 49-65).

84. J. Turnure. Reactions to physical and social distractors by moderately retarded institutionalized children. Research Report #3. June 1970. (Journal of Special Education, 1970, 4, 283-294).

85. D. Moores. Evaluation of preschool programs: An interaction analysis model. Occasional Paper #1. April 1970. (Keynote Address, Diagnostic Pedagogy, International Congress on Deafness. Stockholm, August 1970, also presented at American Instructors of the Deaf Annual Convention, St. Augustine, Florida, April, 1970).

86. D. Feldman & W. Markwalder. Systematic scoring of ranked distractors for the assessment of Piagetian reasoning levels. Research Report #2. March 1970. (Educational and Psychological Measurement, 1971, 31, 347-362).

87. D. Feldman. The fixed-sequence hypothesis: Individual differences in the development of school related spatial reasoning. Research Report #1. March 1970.